

AVb45/55 AVb45/55

FEATURES	1. .769
source	

1. .769

[illegible][illegible]

Query Match	32.38	Score 457.8	EB 9	Length 792
Best Local Similarity	97.78	Pred. No. 1,2e-129		
Matches 505	Conservative 0	Mismatches 12	Indels 0	Gaps 0

```
<urlism "http://www.csis.org">  
/ch_xref="axxon-9606"  
/clone="IMAGE:369033g"  
/clone_lib="NIH_MRG_44"  
/issue_type="endometrium, adenocarcinoma (<<  
/lab_host="PHILIP (Phage-Resistant)"  
/note="Organ: uterus; Vector: pOTB7; Site:1"
```


The first part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present. The second part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present.

F1	Region	56	"potential phosphorylation site"
FT	Region	64	"potential phosphorylation site"
FT	Region	76	"potential phosphorylation site"
FT	Region	185	"potential phosphorylation site"
FT	Region	211	"potential phosphorylation site"
FT	Region	234	"potential phosphorylation site"
FT	Region	245	"potential phosphorylation site"
FT	Region	248	"potential phosphorylation site"
FT	Region	287	"potential phosphorylation site"
FT	Region	289	"potential phosphorylation site"
XX			
FN	W0200028042-A2.		
XX			
PD	18-MAY-2000.		
XX			
PF	12-N-V-1999;	99WO-US26742.	
XX			
PR	12-N-V-1998;	98US-0191280.	
PR	07-DEC-1998;	98US-0206547.	
PR	08-MAR-1999;	99US-0123404.	
XX			
PA	(INCYTE) INCYTE PHARM INC.		
XX			
P1	Tang YT, Corley NC, Guegler KJ, Yue H, Baughn MR, Lai P;		
FI	Hillman JL, Burdman G, Azimciak Y, Au Young J;		
XX			
D8	WPI: 2000-376546/32.		
XX			
PT	New human cell surface receptor protein and polynucleotide useful for		
PT	diagnosis, prevention and treatment of cancer, immune disorders,		
PT	infection and neuronal disorders -		
XX			
PS	Claim 1; Page 86-87; 97pp; English		
XX			
CC	The present sequence is a novel human cell surface receptor protein		
CC	(hSRP) designated hCSR-12. The nucleotide sequence was identified in		
CC	Incyte Clone 3344986 from the CDNA library SPLN0709, which was made		
CC	from RNA isolated from diseased spleen tissue. A number of Incyte Clones		
CC	were used to assemble the consensus sequence. BLAST analysis showed that		
CC	the sequence is homologous to non-CD4 glycoprotein gp120 receptor		
CC	AAC32188. hCSR-12 and its analogs are useful for preventing or treating		
CC	disorders associated with decreased or increased expression or activity		
CC	of hCSR-12. Such disorders include cancers such as leukemia and melanoma,		
CC	immune disorders such as rheumatoid arthritis, asthma and		
CC	atherosclerosis, bacterial and parasitic infections and neurological		
CC	epilepsy. Polynucleotides encoding hCSRs may be used as hybridisation		
CC	probes to diagnose these conditions. Anti-hCSR antibodies may be used		
CC	as antagonists, as a targeting or delivery mechanism for binding		
CC	pharmaceutical agents into contact with cells or tissues expressing		
CC	hCSR-12 and for diagnosis of hCSR-related disorders. hCSR-12 and its		
CC	catalytic or immunogenic fragments are useful for drug screening using		
CC	libraries of compounds.		
XX			
S2	Sequence 325 AA:		
Query Match	100.0%; Score 1696; E= 21; Length 325.		
Best local Similarity	100.0%; Prec. No. 3 2e+140;		
Matches 325.	Conservative 0. Mismatched 0. Indels 0. Gaps 0		
QY	1 MSDSPPPVGGTALTCGCRALVALGIQLSFMILAGYVAIIQQVSVPSSLSVEQSRTCA 60		
1	MSDSPPRVGLDLCRIGHGVLIQLLSPMLIAGVLAIIQVSVSPSSIQESGHQA 60		

27		61	IYGRDQAAVTSERSEGETVEFTQQKAAVTEHPKSKGLGYOETPLKAAGE	120
28		61	IVYNTQQPAAVDELSDFEKUDELDTGKAAVQLPFRSKIQPTVQPTLKAAGF	120
29		121	GTFSTKNGTQYTHLPKSAWAEIPEKSTGEHYQGLHPLKAAVFEPLTPSKAQITVEL	180
30		121	IEEACQELGYEDLPAVAWELEKRSKEIYDGLHPLKAAVTEHPSSKQEQETVEL	180
31		181	HTTAETEEPPDPFWTFWTFFYSNRYMFGSLFMWHISLVADGVFAALVYTFALQNH	240
32		181	IDLKAFERDLRKHCKRWMLFPQGNLYEMSNQRKWHSVIACDEVRAQLVLKTAEQNPF	240
33		241	LQLTSPENPSWMMLSTLNDEGTAWWWSLSPSPPOFYNNSEPNNSGNPCPAPPSGS	300
34		241	EELQTSFNPSWMMLSTLNDEGTAWWWSLSPSPPOFYNNSEPNNSGNPCHAPPSGS	300
35		301	GMNIHQEVWNYWCTEPFAATPEDE	325
36		301	GMNIHQEVWNYWCTEPFAATPEDE	325
RESULT 2				
AAB19714	ID	AAB19714 standard; Protein: 404 AA.		
XX	AC	AAB19714;		
XX	DT	19-FEB-2001 (first entry)		
DE		Dendritic cell specific C-type lectin DC-STGN.		
XX	KW	Dendritic cell; lectin; DC-STGN; intercellular adhesion molecule;		
KW		ICAM-3 receptor; human; HIV-1 infection; autoimmune disease;		
XX		allergy; immunotherapy; immunosuppressive; therapy.		
OS		Homo sapiens.		
XX	PN	EP1046651-A1.		
XX	PD	25-OCT-2000.		
XX	PF	19-APP-1999; 99EP-0201204.		
XX	PR	19-APP-1999; 99EP-0201204.		
XX	PA	(UYNI-) UNIV KOENINK NIMEGSEN.		
P1		Figdor CG, Gelfandbeck THH, van Kooyk Y, Torensmas R;		
XX		Witt: 2000 648929/63.		
XX		N-PSSB; AAA88740.		
PT		Modulating the immune response in an animal, useful e.g. for		
PT		immunosuppression, using a compound that binds to C-type lectin on		
PS		dendritic cells		
PS		Claim 11; page 18-20; 4app; English.		
CC		The present sequence is that of human DC STGN, a 44 kDa C type		
CC		lectin of dendritic cells (DC). DC-STGN is a novel ICAM-3 receptor		
CC		that is specifically expressed by human DC and is involved in the		
CC		initial transient DC-T-cell interaction necessary for initiating an		
CC		immune response. DC-STGN is also able to bind to the HIV envelope		
CC		protein gp120 and to facilitate HIV-1 entry into DC. The invention		
CC		relates to the use of a compound that binds to a C-type lectin		
CC		(especially DC-STGN) on the surface of a DC, and in the preparation		
CC		of a composition for modulating, especially reducing, the immune		
CC		response in an animal, in particular a human or other mammal. The		
CC		composition modulates the interactions between a DC and a T-cell,		
CC		more specifically between a C-type lectin on the surface of a DC		
CC		and an ICAM receptor on the surface of a T cell. It is used to		
CC		prevent/inhibit immune responses to specific antigens, for inducing		

CC tolerance, for immunotherapy, for immunosuppression, for the
 CC treatment of autoimmune diseases, the treatment of allergy, and/or
 CC for inhibiting HIV infection (claimed). The compound that binds to
 CC the C-type lectin is chosen from mannose, fucose, plant lectins,
 CC antibiotics, sugars, proteins and antibodies against C-type
 CC lectins, especially against DC-SIGN. Anti-DC-SIGN antibodies were
 CC developed that not only inhibited transient DC-1-cell interactions
 CC and DC-induced T-cell proliferation but also inhibited HIV-1
 CC infection of DCs.

CC Sequence 404 AA:

Query Match 83.3% Score 1412.5; DB 21, Length 404;
 Best Local Similarity 71.9%; Pred. No. 5,60-107;
 Matches 282; Conservative 20; Mismatches 19; Indels 71; Gaps 4;

QY 1 MSDSKEPRVQGLGL-----OCIGHGALVGLLSPMLASVLA 40
 DB 1 MSDSKEPRVQGLGLLEEQRLGIFRQYKSLAGCLGHPLVQLLSEFTLAG----L 56
 QY 41 LVQVSKVPSLSIQSESDAIVQNLITLKAANGELSEKSLQEIYQELITLKAANGELPE 100
 DB 57 LVQVSKVPSLSIQSESDAIVQNLITLKAANGELSEKSLQEIYQELITLKAANGELPE 116
 QY 101 KSKLQEIYQELITLKAANGELPEKSKL----- 127
 DB 117 KSKLQEIYQELITLKAANGELPEKSKLQEIYQELITLKAANGELPEKSKMQEIYQELITL 176
 QY 128 -----QEIYQELITLKAANGELPEKSKLQEIYQELITLKAANGELPEKSKQ 174
 DB 177 KAANGELPEKSKQEIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQ 236
 QY 175 QIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQ 236
 DB 237 QIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQ 296
 QY 235 AEFQNLQIQTSSNPFSSMGLSDINQPGTWQVWVQSPISFQPYWNSGCEPNNSGNEQC 294
 DB 297 AEFQNLQIQTSSNPFSSMGLSDINQPGTWQVWVQSPISFQPYWNSGCEPNNSGNEQC 355
 QY 295 AEFSSGMMNDKQNDVNWICKKPA--CFEPE 325
 DB 457 AEFSSGMMNDKQNDVNWICKKPA--CFEPE 388

RESULT 3

AAB28614 ID AAB28614 standard; Protein: 404 AA.

AC AAB28614:

DT 12-FEB-2001 (first entry)
 XX Human C-type lectin receptor.
 DE Human C-type lectin receptor.
 XX Human: immunomodulatory; antiallergic; anti-HIV; antiinflammatory;
 KM antidiabetic; antihypertoid; antirheumatic; antiallergic; vaccine;
 KM C-type lectin receptor; immunotherapy; immunosuppression;
 KM transplant rejection; autoimmune disease; thyroiditis;
 KM rheumatoid arthritis; multiple sclerosis; autoimmune diabetes;
 KM systemic lupus erythematosus; HIV infection; allergy.
 XX Homo sapiens.

OS Homo sapiens.

PN W020006.1251-A1.

PD 26-OCT-2000.

PF 19-APR-2000; 2000MO-NL00253.

PK 19-APR-1999; 99EP-0201204.

PR 20-JAN-2000; 200005-0176924.

XX

PA (UWI-) UNIV NIJMEGEN.

PI Figdor CG, Geijtenbeek TH, Van Kooyk Y, Forstma R

XX WPI: 2000-656424/63.

DR N-PSDB: AAC65383.

XX Methods for modulating dendritic cell and T cell interaction
 PT compound, such as mannose carbohydrates, that binds to a C-type
 PT on a dendritic cell surface, useful in the treatment of auto-
 PT diseases and allergy -

PS Disclosure: Fig 9; 66pp; English.

CC The present sequence is DC-SIGN, a HLA-C type lectin receptor
 CC methods for modulating dendritic cell and T cell interaction
 CC disclosed. The methods involve using a compound that binds to
 CC lectin on the surface of a dendritic cell. The methods are use-
 CC preventing or inhibiting immune responses to specific antigens
 CC inducing tolerance, for immunotherapy, for immunosuppression (e.g.
 CC preventing transplant rejection), for the treatment of autoim-
 CC diseases (e.g., thyroiditis, rheumatoid arthritis, multiple sys-
 CC autoimmune diabetes, systemic lupus erythematosus), HIV infec-
 CC and/or for the treatment of allergy. Antibodies are useful for
 CC detecting the presence of dendritic cells in a biological sam-
 CC for determining the presence and/or expression of the C-type
 CC or their fragments or epitopes in a biological sample.

XX Sequence 404 AA:

Query Match 83.3% Score 1412.5; DB 21, Length 404;
 Best Local Similarity 71.9%; Pred. No. 5,60-107;
 Matches 282; Conservative 20; Mismatches 19; Indels 71;

QY 1 MSDSKEPRVQGLGL-----OCIGHGALVGLLSPMLASVLA 40
 DB 1 MSDSKEPRVQGLGLLEEQRLGIFRQYKSLAGCLGHPLVQLLSEFTLAG----L 56
 QY 41 LVQVSKVPSLSIQSESDAIVQNLITLKAANGELSEKSLQEIYQELITLKAANGELPE 100
 DB 57 LVQVSKVPSLSIQSESDAIVQNLITLKAANGELSEKSLQEIYQELITLKAANGELPE 116
 QY 101 KSKLQEIYQELITLKAANGELPEKSKL----- 127
 DB 117 KSKLQEIYQELITLKAANGELPEKSKLQEIYQELITLKAANGELPEKSKMQEIYQELITL 176
 QY 128 -----QEIYQELITLKAANGELPEKSKLQEIYQELITLKAANGELPEKSKQ 174
 DB 177 KAANGELPEKSKQEIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQ 236
 QY 175 QIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQ 236
 DB 237 QIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQEIYQELITLKAANGELPEKSKQ 296
 QY 235 AEFQNLQIQTSSNPFSSMGLSDINQPGTWQVWVQSPISFQPYWNSGCEPNNSGNEQC 294
 DB 297 AEFQNLQIQTSSNPFSSMGLSDINQPGTWQVWVQSPISFQPYWNSGCEPNNSGNEQC 355
 QY 295 AEFSSGMMNDKQNDVNWICKKPA--CFEPE 325
 DB 357 AEFSSGMMNDKQNDVNWICKKPA--CFEPE 388

RESULT 4

AAG79086 ID AAG79086 standard; Protein: 404 AA.

AC AAG79086:

DT 10-DEC-2001 (first entry)
 XX Human DC-SIGN, a dendritic cell-specific C-type lectin.
 DE Human DC-SIGN, a dendritic cell-specific C-type lectin.
 XX


```

QY 1 MSDSKEPRVQOGLL-----GTLGHGALVLDLSFMLAGLVAI 40
DB 1 MSDSKEPRVQOGLLLEFQRLPGIFPRQTPGYKSLAGVLCRHGLVLDLISFTLLAG-----L 56
QY 41 LVQVSKVPSSLSQBSQSDAIYQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 100
DB 57 LVQVSKVPSSLSQBSQSDAIYQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 116
QY 101 KSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 127
DB 117 KSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 176
QY 128 -----QELYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 174
DB 177 KAAVGLPEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 246
QY 175 QLYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 244
DB 247 ELYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 296
QY 245 AEFQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 294
DB 297 AEFQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 346
QY 295 AEFQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 325
DB 457 AEFQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 388

```

RESULT 6

AAK32188
ID AAK32188 standard; Protein; 404 AA.

AAK32188;

08-JUN-1993 (first entry)

Sequence of a non-CD4 glycoprotein gp120 receptor protein.

Surface cell protein; glycoprotein receptor protein; gp120 receptor;

HIV; diagnosis; therapy.

Homo sapiens.

W09301820-A.

04-FEB-1993.

16-JUN-1992: 92MO-0505985.

16-JUL-1991: 91US-0731214.

(BRIM) BRISTOL-MYERS SQUIBB CO.

Curtis BM;

WPI: 1993-058515/07.

DR N-PSDB: AAK36648.

Inhibition of non-CD4 mediated HIV infection - dependent on novel

gp120 receptor, for diagnosing and treating HIV infection

Disclosure; Figure 3A; 44pp; English.

The gp120 receptor protein is present on placental, muscle, neural, brain, dendritic or mucosal cells. It may be produced recombinantly by cloning cDNA isolated from a library of recombinant placental genes. Binding of gp120 to its receptor is inhibited by specific carbohydrates, plant lectins and specific antibodies. It corresponds to a non-CD4 mammalian cell surface protein, contg. ca. 400 AAs and having a mol. wt. of ca. 45 kD, and a binding affinity for gp 120 of Kd = 1.32-2 nM.

XX
SQ Sequence 404 AA;

Query Match 82.3%; Score 1396.5; DB 14; Length 4
Best Local Similarity 71.4%; Prog. No. 110-105;
Matches 280; Conservative 21; Mismatches 20; Indels 7

```

QY 1 MSDSKEPRVQOGLL-----GTLGHGALVLDLSFMLAGLVAI 40
DB 1 MSDSKEPRVQOGLLLEFQRLPGIFPRQTPGYKSLAGVLCRHGLVLDLISFTLLAG-----L 56
QY 41 LVQVSKVPSSLSQBSQSDAIYQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 100
DB 57 LVQVSKVPSSLSQBSQSDAIYQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 116
QY 101 KSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 127
DB 117 KSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 176
QY 128 -----QELYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 174
DB 177 KAAVGLPEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 246
QY 175 QLYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 244
DB 247 ELYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 296
QY 245 AEFQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 294
DB 297 AEFQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 346
QY 295 AEFQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 325
DB 357 AEFQNTLTQKAAVGLSEKSKLOEYQELTLQKAAVGLSEKSKLOEYQELTLQKAAVGLPE 388

```

RESULT 7

AAW88493
ID AAW88493 standard; Protein; 296 AA.

AAW88493;

30-MAR-1999 (first entry)

Human liver clone HP01347; human; lectin; receptor; liver.

Transmembrane protein; HP01347; human; lectin; receptor; liver.

Homo sapiens.

W09855508-A2.

10-DEC-1998.

03-JUN-1998: 98MO-TP02445.

03-JUN-1997: 97JP-0144948.

(PROT-) PROTESSE INC.

(SAGA) SAGAMI CHEM RES CENTRE.

Kato S, Sekine S, Yamaguchi T;

WPI: 1999-045730/04.

DR N-PSDB: AAW84361.

New human proteins containing transmembrane domains and their encoding sequences - useful in the preparation of antibodies at large-scale protein production, gene diagnosis, and gene therapy.

Claim 4; Page 125-126; 178pp; English.

This is the amino acid sequence of a transmembrane protein encoded

Sequence	296 AA:
50	

Query Match	Score	DB	Length
68.28;	1157;	20;	2967;

Best Local Similarity	83.84;	Field No.	2111-84;	Indels	46;	Gaps	1
Matches	238;	Conservative	0;	Mismatches	0;		

1

5.

Sequence	AA;
54	

Query Match	Score	DB 22	Length
68.28	1157		296

Rest. Local Similarity	R3 R4:	Pred. NO. 2.0e-09;	Indels	45;	Subs	1,
Rest. Conservative	0:	Mismatches	0,			

Matches 23

1 MSDSKPRVQQLGCLGHGALVQLLSFMLLAGVLAIVQVSNVPSQUSQLEGGN

Db 1 MSDSKPRVQQLGLGCIHGALVLQLSFMLLAGVLVAILVQVSKVPSSLSQEQSEQDA 60

12 KAVGEITRKYOEIPEKSKLOEIKAVGELTOKIYOEISEKSGVETKAVGE

[illegible]

61 IYQNTQLKAAVGELSEKSLQEIYQELTQLKAAVGELPEKSLQEIYQELINENNAVG

121 IPEKSKI,QEIYQEI,TRI,KAVGELPEKSI,QEIYQELT-----1

18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
8

Dd 121 LPEKSKLQEIYQFIKLAAVGEIENOVGZAK

QY 160 -----LKAAGELPDQSKQQLYQELIDLNIAFENENKNS

[18] TEIKAVGELPEKSKLOH IYQEL.TQI.KAVGELPDÖSKÖQ IYQELTDKTAFFELCRHC 24

[illegible]

QY	195	PRDWLEFQUNCYFMSNSQNWMDSVLWQOYVWQZT	195
QY	195	PRDWLEFQUNCYFMSNSQNWMDSVLWQOYVWQZT	195

Db 241 PKDWTFQGNCFMNSQRNHDSVTACQEVRAQLVIKTAEQ 284

RESULT 9

ABB09715
ID ABB09715 standard; Protein; 296 AA.

ADDDOYAC
ID
XX

AC ABB09715;
yy

AA
DT 11 JUN-2002 (first entry)

XX amino acid sequence of human polypeptide HP01347.

DE
XX

domain protein: HP01347.

Antibody; antigen; transmembrane domain; protein; ...

AA
OS Homo sapiens.

XX
XX
W0900308416-A1

W0200000410 C11.
PN
XX
XX

PD 31-JAN-2002.

XX
PF 24-JUL.-2001: 2001WO-JP06371.

XX 2000-2000 1P-03227A3

ER 24-JUL-2000; 2000JP-0254407.
ER 24-AUG 2000; 2000JP-0254407.

XX
XX
TECHNOLOGY CORP.

PA (NISC-) JAPAN SUI & TECHNOLOGY CORP.
XX
XX

PI Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K, Ishizuka Y,
XX WPI: 2002-195877/25.
DR N-PSDB: ABK41993.

PT Antibody preparation by inoculation of an animal with a vector
XX expressing a fusion protein of an antigen on the C-terminal side of a
PT transmembrane domain for use as drugs, diagnostic reagents and
PT laboratory reagents

XX Example: Page 23-27; 45pp; Japanese

XX The specification describes a method of antibody preparation. The
CC method comprises inoculating an animal with a vector expressing
CC a fusion protein having an antigen protein fused to the C-terminal
CC side (extracellular) of a transmembrane domain protein (the
CC N-terminal side of which is intracellular), and then isolating and
CC purifying the antibody from the animal. The antibodies can be used
CC as drugs, vaccines, diagnostic reagents and laboratory reagents. The
CC present sequence represents a polypeptide, designated HP01347, which was
CC used in the course of the invention.

XX Sequence 296 AA:

Query Match 68.2% Score 1157; DR 23; Length 296;
Best Local Similarity 84.8%; Pred. No. 2; 6e-86;
Matches 238; Conservative 0; Mismatches 0; Indels 46; Gaps 1;

OY 1 MSDSKPRVQQLGLLGLGHNALVIGLSPMLLAGVIAIIVGVSKVPSLSQESQCA 60
DB 1 MSDSKPRVQQLGLLGLGHNALVIGLSPMLLAGVIAIIVGVSKVPSLSQESQCA 60
OY 61 IVGNLTQAAVSELSEKSLQEIYQELTQAAVGELEPKSKLOEYQELTRKAWE 20
DB 61 IVGNLTQAAVSELSEKSLQEIYQELTQAAVGELEPKSKLOEYQELTRKAWE 20
OY 121 LPEKSKLOEYQELTRKAAGVLEPKSKLOEYQELTRKAAGVLEPKSKLOEYQEL 159
DB 121 LPEKSKLOEYQELTRKAAGVLEPKSKLOEYQELTRKAAGVLEPKSKLOEYQEL 180
OY 160 -----LKAAGVLEPKSKLOEYQELTQAAVGELEPKSKLOEYQELTRKA 194
DB 161 LELKAAGVLEPKSKLOEYQELTQAAVGELEPKSKLOEYQELTQAAVLEPKSK 240
OY 195 PKQWTFQGMCTFMSNSQPMHDSVTACQEPFAOTVYTKFAEQ 238
DB 241 PKQWTFQGMCTFMSNSQPMHDSVTACQEPFAOTVYTKFAEQ 284

RESULT 10
AA088024
ID AA088024 standard; Protein: 325 AA.

XX AA088024:

DT 05-JUN-2002 (first entry)

DE Mouse protein encoded by Contig 1A.

XX Nucleic acid library; immune response; asthma; COPD;
KM airway hyperresponsiveness; bronchoalveolar manifestation;
KM signature sequence; SS; chronic obstructive pulmonary disease;
KM allergic disease; rhinitis; atopic dermatitis; urticaria;
KM autoimmune disease; multiple sclerosis; inflammatory bowel disease;
KM allograft rejection; infectious disease;
KM calcium-activated chloride channel.

XX Mus sp.

XX W0200214366-A2.

XX 21-FEB-2002.

PF 16-AUG-2001; 2001W0-ML00610
XX 16-AUG-2001; 2001W0-ML00610
XX 16-AUG-2001; 2001W0-ML00610
XX 16-AUG-2001; 2001W0-ML00610

XX (OYOT-) RIKSUNIV UTRECHT.

XX Grool PC, Van Berghenhouwen HJ, Van Gosterhout AJM;
XX WPI: 2002-241888/29.

XX N-PSDB: ABK41992.

XX Nucleic acid library comprising genes which are capable of
PT progression and suppression of an immune response, especially
PT response observed with airway hyper-responsiveness of asthma

XX Example 10; Fig 10; 120pp; English.

XX The invention relates to a nucleic acid library comprising ge
CC their fragments which are capable of modulating an immune res
CC observed with airway hyper-responsiveness and/or bronchoalveo
CC manifestations of asthma. Also included are a method for modu
CC immune response of an individual comprising modulating a gene
CC a nucleic acid at least functionally equivalent to a nucleic
CC as R1-S0-R1-A11, S001-A10, S002-1-C11, S001-A12, and R1-S0-R1
CC substance (for use as a medicament) capable of modulating a
CC comprising a nucleic acid at least functionally equivalent to
CC acid identifiable by SS and the use of a proteinaceous subst
CC from a nucleic acid at least functionally equivalent to a nu
CC identifiable by SS for the production of an antagonist (for us
CC medicament) against the substance. The antagonist and substanc
CC useful for the treatment of an immune response observed with
CC hyperresponsiveness and/or bronchoalveolar manifestations of
CC The method is useful for modulating the above immune respons
CC gene encodes a gene product capable of modulating the immune
CC The substance is useful for treating an immune response, part
CC asthma, chronic obstructive pulmonary disease (COPD), allergic
CC (rhinitis atopic dermatitis, urticaria), autoimmune diseases
CC multiple sclerosis), inflammatory bowel disease, allograft re
CC infectious disease. The present sequence is a mouse or human
CC protein encoded by a signature sequence gene or its homologue
CC equivalent.

XX Sequence 325 AA:

Query Match 42.5%; Score 720; DR 23; Length 325;
Best Local Similarity 43.0%; Pred. No. 1; 3e-50;
Matches 154; Conservative 49; Mismatches 77; Indels 7;

OY 1 MSDSKPRVQQLGLLGLGHNALVIGLSPMLLAGVIAIIVGVSKVPSLSQESQCA 60
DB 1 MSSTKAKMQLSSMDDELWVSGSPYSTPSPSPSPSPSPSPSPSPSPSPSPSPSP 60
OY 30 FMLLAGVIAIIVGVSKVPSLSQESQCAIYQNLITL-----KAAGVLEPKSK 60
DB 61 FFLAGVIAIIVGVSKVPSLSQESQCAIYQNLITL-----KAAGVLEPKSK 60
OY 84 IYQELTQAAVSELSEKSLQEIYQELTQAAVGELEPKSKLOEYQELTRKAWE 20
DB 119 IYQELTQAAVSELSEKSLQEIYQELTQAAVGELEPKSKLOEYQELTRKAWE 20
OY 144 LPEKSKLOEYQELTRKAAGVLEPKSKLOEYQELTRKAAGVLEPKSKLOEYQEL 159
DB 158 -----ELSKISFPVYDSDKQKERTYQQLVQMKTELFRTRK 240
OY 202 QGNCYFMSNSQPMHDSVTACQEPFAOTVYTKFAEQ 238
DB 203 LGNCYFMSNSQPMHDSVTACQEPFAOTVYTKFAEQ 284
OY 262 EGIMQWTFQGMCTFMSNSQPMHDSVTACQEPFAOTVYTKFAEQ 325
DB 262 EGIMQWTFQGMCTFMSNSQPMHDSVTACQEPFAOTVYTKFAEQ 325

anti-bacterial; endocrine; cardiac; central nervous system; virulence;
anti-HIV; fungicide; anti-invasive; cardiovascular; anti-nausea;
anti-asthmatic; haemostatic; pulmonary; anti-cancer; osteoporosis; eczema;
dermatological; anti-allergic; anti-asthmatic; anti-infective; cytostatic;
neuroprotective; antidepressant; neurotrophic; antiparkinsonian; infection;
immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
anti-anaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
cardiac dysfunction; neuropathology; cardiac anaphylaxis; autografting;
genetic disease; haematopoietic disorder; platelet disorder; asthma;
thrombocytopenia; osteoporosis; septic shock; multiple sclerosis;
dermatological; diabetes; multiple sclerosis; depression;
Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
neurological disorders.

Sequence 143 AA:

Query Match: 31.18, Score 527, DB 22, Length 143,
Best Local Similarity 78.24, Prod. No. 2.5e-15;
Matches 115; Conservative 4; Mismatches 4; Indels 24; Gaps 2;

1 43 QVSKVPSSTSSQSSQDAIVQNTTQAKAVGPISEKSKLOEIVQELTQAKAVGDELPERS 102
117 57 QVSKVPSSTSSQSSQDAIVQNTTQAKAVGPISEKSKLOEIVQELTQAKAVGDELPERS 116

103 K12ETVAGLTPVKAAGVLPKPSYQGE 129
117 K12ETVAGLTPVKAAGVLPKPSYQGE 143

RESULT 15
ID ABB72372 standard; Protein; 193 AA.
AC ABB72372;
DE 04-APR-2002 (first entry)

Murine protein isolated from skin cells SEQ ID NO: 696.

Human, rat, mouse, skin cells, skin wound, cancer, growth defect;
developmental defect; inflammatory disease; dermatological; pulmonary;
immunomodulator, anti-inflammatory; cytostatic; neuroprotective.

Mus sp.
WO200190357-A1.
29 NOV 2001.

24-MAY-2001, 2001W0 H200099.
24-MAY-2000, 2000US 206650P.
25-JUL-2000, 2000US 22133P.

(GENE-) GENESIS RES & DEV CORP LTD.

Watson JF, Strachan T, Sleeman M, Varust K, Mullison JB, Kahlke KB,
N-PSDB, AB050472.

How polynucleotides and polypeptides encoded by the polynucleotides
isolated from skin cells, useful for treating skin wounds, cancers,
growth and developmental defects, inflammatory diseases, or for
modulating immune responses

Claim 4, Page 447 446; 466pp; English.

The present invention provides the protein and coding sequences of GNAS
isolated from human, murine and rat skin cell libraries. The sequences
can be used in the development of therapeutic agents useful in the
treatment of skin diseases, including skin wounds, cancer, growth
defects, developmental defects and inflammatory diseases. The proteins
have important roles in the induction of hair growth, cell proliferation
and cell-cell interaction, in maintaining tissue integrity. In wound
healing and in modulating immune responses, the present sequence is a
polypeptide of the invention.

Sequence 193 AA:

Query Match: 24.58, Score 415, DB 24, Length 193,
Best Local Similarity 49.84, Prod. No. 5e-26;
Matches 74; Conservative 34; Mismatches 55; Indels 24; Gaps 3;

156 EITPLKA-AVGHLPDQ-----SKQQIVQELTDLDTAFPRDPRHC 194
6 QVSKVPSSTSSQSSQDAIVQNTTQAKAVGPISEKSKLOEIVQELTQAKAVGDELPERS 65

195 PRQPTFGANVEMNSQPMWHSVTAGVPAQVAVIKTAEDBNFLDTSPNSPSSMM 254
66 PMWELTQSSVYLSKLTGSMETSASSPDELGAHLIVNVSDBPRPKYNNVPRNOSMI 125

255 G1S0LNGGQWQVWVNSPISPPQRYNNSCPNNSNENPCAFPSGQWNNRCVNDVW1 314
126 G1S0LTHGSGWQWVNSALRFSF---WKESPPNNDKEDCVLELPMDDWNNKCTEONWV 182

Tue Dec 10 10:19:07 2002

us-09-831-458a-12.std.rag

Page 11

OY 315 CKKPAA 320
I::I::I
DB 183 GEOPSA 188

Search completed: December 7, 2002, 11:03.57
Job time : 76 secs

•
•
•
•

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuLink

OM protein - protein search, using SW model

Run on: December 7, 2002, 10:41:06 : Search time 44 seconds
(without alignments)
710,084 Million cell updates/sec

Title: US-09-831-458A-12
Perfect score: 166
Sequence: 1 MSDSKEPRVQQLLEDELCH RGVNRWVCKKAAVHLE 325

Scoring table: RUSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 9614422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: 1: PIR_73:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1412.5	83.3	404	2 A46274	HIV gp120-binding
2	351	20.7	309	1 S34198	IgE Fc receptor II
3	350	20.6	331	1 LNMSER	IgE Fc receptor, I
4	326.5	19.3	207	1 LNCHIL	hepatic lectin - C
5	322.5	19.0	321	1 LNHRFP	IgE Fc receptor II
6	309.5	18.2	550	2 A28165	Kupffer cell recep
7	307	18.1	304	2 JX0209	lectin, galactose/
8	306	18.0	284	2 S29855	asialoglycoprotein
9	289	17.0	284	1 LNRTL	hepatic lectin - r
10	287	16.9	306	2 A42230	lectin M-ASGP-BP p
11	271.5	16.0	291	1 LNHD1	hepatic lectin HI
12	267	15.7	311	1 LNHR2A	asialoglycoprotein
13	258	15.2	301	2 S13165	asialoglycoprotein
14	243	14.3	301	2 LNRT2	hepatic lectin 2 -
15	232	13.7	237	2 J07608	type II lectin-11k
16	230.5	13.6	363	2 J07608	lectin-like oxidiz
17	230	13.6	742	2 J07595	scavenger receptor
18	227.5	13.4	166	2 J28141	C-type lectin, B 1
19	219.5	12.9	162	1 LNRC3	lectin BRA3-1 prec
20	218.5	12.9	162	1 LNRC1	lectin BRA3-2 prec
21	216.5	12.8	216	2 P07375	natural killer cel
22	210.5	12.4	1456	1 A46563	mannose receptor p
23	203.5	12.0	330	2 T46256	brevican - human
24	202.5	11.9	223	1 A75917	NK-cell receptor p
25	202.5	11.9	2415	1 A46086	aggrexin precursor
26	201.5	11.9	2124	2 A28452	aggrexin core
27	199.5	11.8	1479	2 T42710	mannose receptor,
28	199	11.7	883	2 S57653	brevican precursor
29	198	11.7	257	2 I50146	gene 17.5 protein

30	197	11.6	2132	1 A55182	aggrexin
31	195	11.5	883	2 S49126	brevican
32	195	11.5	2109	1 I50421	aggrexin
33	194.5	11.5	912	2 A54423	brevican
34	193.5	11.4	1340	2 A39608	aggrexin
35	193.5	11.4	2327	2 T42630	aggrexin
36	191	11.3	199	2 JH0822	lymphocyte
37	189.5	11.2	1268	2 S52781	neutrophil
38	188	11.2	237	2 A44467	neutrophil
39	188	11.1	1455	1 A48925	mannose
40	186.5	11.0	1257	1 S28764	neutrophil
41	180.5	10.6	147	2 A26697	echinoid
42	178	10.5	244	1 LNMSMC	mannose
43	176.5	10.4	3562	2 A47171	chondroitin
44	175	10.3	412	2 P42755	E-selectin
45	174.5	10.3	391	2 T34284	hypoderm

ALIGNMENTS

RESULT 1

A46274
HIV gp120-binding C-type lectin - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 01
C:Accession: A46274
C:Curator: B.M.; Scharnowski, S.; Watson, A.J.
Proc. Natl. Acad. Sci. U.S.A. 89, 8156-8160, 1992
A:Title: Sequence and expression of a membrane-associated C-type
A:Reference number: A46274; MUID:92490446; PMID:1518665
A:Accession: A46274
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-404 <gap>
A:Experimental source: placenta
A:Note: sequence extracted from NCBI: loc:loc (NCBI:113134, NCBI)
C:Superfamily: C-type lectin homology
F:256-377/Domain: C-type lectin homology <LCH>

Query Match 83.3% Score 1412.5 DB 2: Length 404
Best Local Similarity 71.9% Pred. No. 126-78;
Matches 282; Conservative 20; Mismatches 19; Indels 7

QY	1	MSDSKEPRVQQLLEDELCH	-----GCTGGALVQLLSFPMGLAT
DB	1	MSDSKEPRVQQLLEDELCH	-----GCTGGALVQLLSFPMGLAT
QY	41	LVQYSKVPSSLSQSESDAIIYQNTLQI	-----GCTGGALVQLLSFPMGLAT
DB	57	LVQYSKVPSSLSQSESDAIIYQNTLQI	-----GCTGGALVQLLSFPMGLAT
QY	101	KSKEDEIYQELTRKAAGVLEPEKSK	-----GCTGGALVQLLSFPMGLAT
DB	117	KSKEDEIYQELTRKAAGVLEPEKSK	-----GCTGGALVQLLSFPMGLAT
QY	128	-----EFYQELTFYAAVETPEVSK	-----GCTGGALVQLLSFPMGLAT
DB	177	KAAGVLEPEKSKQEIYQELTRKAAGV	-----GCTGGALVQLLSFPMGLAT
QY	175	QIYQELTDKIAVEKLECHQETFTW	-----GCTGGALVQLLSFPMGLAT
DB	237	EIYQELTDKIAVEKLECHQETFTW	-----GCTGGALVQLLSFPMGLAT
QY	235	AEQNFLOLSQSPNSPMSMGLSDI	-----GCTGGALVQLLSFPMGLAT
DB	237	EIYQELTDKIAVEKLECHQETFTW	-----GCTGGALVQLLSFPMGLAT
QY	237	AEQNFLOLSQSPNSPMSMGLSDI	-----GCTGGALVQLLSFPMGLAT
DB	237	EIYQELTDKIAVEKLECHQETFTW	-----GCTGGALVQLLSFPMGLAT
QY	295	AETSSGKMGNGVDVNTWICKKPA	-----GCTGGALVQLLSFPMGLAT
DB	357	AETSSGKMGNGVDVNTWICKKPA	-----GCTGGALVQLLSFPMGLAT

F146-321/Domain: extracellular #status: predicted -EXT-
 F164-84/Region: 21-residue repeat
 F181-321/Region: 21-residue repeat
 F185-105/Region: 21-residue repeat
 F102-321/Region: 21-residue repeat
 F106-126/Region: 21-residue repeat
 F125-321/Region: 21-residue repeat
 F148-321/Region: 21-residue repeat
 F150-321/Region: 21-residue repeat
 F163-282/Domain: C-type lectin homology -LCH-
 F163-282/Domain: C-type lectin homology -LCH-
 F163-282/Domain: C-type lectin homology -LCH-
 F147-148/Cleavage site: Lys-Leu (unidentified proteinase) #status: experimental
 F149-150/Cleavage site: Arg-Met (unidentified proteinase) #status: experimental
 F191-262-259-273/Cysteine bonds: #status: experimental

Query Match 19.08; Score 322.5; DB 1; Length 321,
 Best local similarity 28.68; Pred. No. 1-10-12;
 Matches 95; Conservative 61; Mismatches 121; Indels 55; Gaps 13;

QY 1 MSDSRKPVQGLDGLGHA--IVQLDLPMLAGVALLVQVSKVSSLSQ--EUSE 57
 DB 1 MRRGQVSEIEIPRRRCRRGTQIVAGIVTAIWALITLILIMHTQTSKQIFRPA 60
 QY 58 QDAIVNTDQKAAG--EISEKSKQATVETLQIKAAVPEPEKSKIQETLPLKA 116
 DB 61 APVSVSKNPFHHRCQMQSKSTQTSPELELVA-----EGQELK 124
 QY 117 AVGELEPKSKLOEYELTETLKAAGCELEPKSK---LAFIYQFIPELKAAGELQVSK 172
 DB 105 Q--DLESLNNLQGLDLSFSKQ--ELNERNSEDLLELREEVTKLR----- 149
 QY 173 QQQYELTDLKATFEFPHICETWTFQJNCYFNSNENHNDVTAQVAVAGV 212
 DB 150 -----MELVSSGCVQCTCPKWNINFCPKCYFEGTKQWHAHYARMEQGVSI 201
 QY 233 KTAEEQNFLOIQTSPNSPFWMGISDINQRTGMQVNGSPGLSPFYNNQSP--NNSGN 294
 DB 202 HSEFQEFILKIASHIG--SWIGKSLILKGLIHWKSHVVS---NMAHGEPEKSKQ 256
 QY 252 EHCAPFSNCTWNNPCTVW NWTGKRTAAQ 321
 DB 257 EDCVMHSGRMADAFDRKLAGWVDEPLATG 288
 RESULT 6
 A28166
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 05-Nov-1999
 C:Accession: A38674; A28166
 F:History: G.W.: H111, P.L.
 J. Biol. Chem. 265, 1850-1857, 1991
 A:Title: Structure of the gene for a carbohydrate-binding receptor analog for rat Kupffer
 A:Reference number: A38674; M01D:91107889; PMID:1846367
 A:Accession: A38674
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-550 <H02>
 A:Cross-references: GH:M55532; NID:420302; PUBM:AAAA0892.1; PID:q203064
 F:History: G.W.: H111, R.L.
 J. Biol. Chem. 263, 7487-7492, 1988
 A:Title: Molecular cloning and sequencing of a cDNA for a carbohydrate binding receptor
 A:Reference number: A28166; M01D:88273433; PMID:306487
 A:Accession: A28166
 A:Molecule type: mRNA
 A:Residues: 1-550 <H02>
 A:Cross-references: GH:J03734; NID:4207050; PUBM:AAA1472.1; PID:q205051
 C:Superfamily: C-type lectin homology
 C:Keywords: transmembrane protein
 F1412-536/Domain: C-type lectin homology -LCH-

Query Match 18.2%; Score 309.5; DB 2; Length 550;
 Best local similarity 28.7%; Pred. No. 1-2e 11,

Matches 81; Conservative 45; Mismatches 110; Indels 45; Gaps 11;

QY 52 EQLGSEPAAYGN--LGLAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 110
 DB 286 AGQGNAMN--LGLIYGLDGLKAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 349
 QY 111 EELKAAVTLDESKKQIYGLDGLKAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 161
 DB 340 LSVNSAL-----KSNVGMKSNIGKAAVYSLKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 394
 QY 162 AAVCHLQVQVYGLDGLKAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 221
 DB 394 AVAHILQVQVYGLDGLKAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 449
 QY 222 QVVPACIWLITAPPGNITGLDGLKAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 280
 DB 440 QVSGAHIAVTSQKAPVQITNAVIG--WGLDGLKAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 497
 QY 281 WNSGEENN--SGN---EDCAEFGSGGWNDRGVDNWDICK 317
 DB 138 WNSGEENN--SGN---EDCAEFGSGGWNDRGVDNWDICK 538
 RESULT 7
 JX0209
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
 C:Accession: JX0209; PX0009
 F:History: M.: Kawakami, K.: Osawa, T.; Toyoshima, S.
 J. Biochem. 111, 331-336, 1992
 A:Title: Molecular cloning and expression of cDNA encoding a galactose/N-acetylglact
 A:Reference number: JX0209; M01D:9248023; PMID:1587794
 A:Accession: JX0209
 A:Molecule type: mRNA
 A:Residues: 1-394 <SNT>
 A:Cross-references: JE:536676; NID:443266; PUBM:AB22171.1; PID:q49361
 F:History: S.: Sato, M.; Toyoshima, S.; Osawa, T.
 J. Biochem. 104, 600-605, 1988
 A:Title: Purification and characterization of a lectin-like molecule specific for gal
 A:Reference number: PX0009; M01D:8917895; PMID:3241002
 A:Accession: PX0009
 A:Molecule type: protein
 A:Residues: 102-120/137; X:149-151 <DA>
 C:Superfamily: hepatic lectin; C-type lectin homology
 C:Keywords: glycoprotein, lectin, macrophage, transmembrane protein
 F136-61/Domain: transmembrane #status: predicted -CPA>
 F173-294/Domain: C-type lectin homology -LCH>
 F174-166/Region: 21-residue repeat (Asn) (covalent) #status: predicted

Query Match 18.18; Score 307; DB 2; Length 304;
 Best local similarity 30.38; Pred. No. 8-7e 12;
 Matches 97; Conservative 59; Mismatches 101; Indels 59; Gaps 16;

QY 27 LLSMLAGVVAIVGVVPSGSPQSGFATVNTDQKAAGPELPSKQIETLPLKA 86
 DB 36 LLSMLAGVVAIVGVVPSGSPQSGFATVNTDQKAAGPELPSKQIETLPLKA 86
 QY 87 ETLQKAAVTLDESKKQIYGLDGLKAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 144
 DB 85 SL-----LSDVSEFEKSTISLKYVD--HROELQGRDLSQVTSISFV-- 128
 QY 134 LLSMLAGVVAIVGVVPSGSPQSGFATVNTDQKAAGPELPSKQIETLPLKA 197
 DB 129 LLSMLAGVVAIVGVVPSGSPQSGFATVNTDQKAAGPELPSKQIETLPLKA 176
 QY 198 WTLQKAAVTLDESKKQIYGLDGLKAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 257
 DB 177 WTLQKAAVTLDESKKQIYGLDGLKAAVDEPSEKTKGLIYGLDGLKAAVDEPEKSKLQGLYOR 234
 QY 258 LLSMLAGVVAIVGVVPSGSPQSGFATVNTDQKAAGPELPSKQIETLPLKA 310
 DB 235 LLSMLAGVVAIVGVVPSGSPQSGFATVNTDQKAAGPELPSKQIETLPLKA 291

F:1-81,87-311/Product: asialoglycoprotein receptor H2b #status predicted -MNT2
 F:1-58/Domain: intracellular #status predicted -INT2
 F:1-23,44-81,87-311/Product: asialoglycoprotein receptor H2c #status predicted -MNT3
 F:59-78/Domain: transmembrane #status predicted -TM2
 F:79-311/Domain: extracellular #status predicted -EXT2
 F:177-300/Domain: C-type lectin homology <LCH>
 F:102,170,305/Binding site: carbohydrate (Asn) (covalent) #status predicted -

Query Match 15.74; Score 267; DB 1; Length 311.

Best Local Similarity 27.24; Pred. No. 2,48-093;

Matches 89; Conservative 49; Mismatches 103; Indels 86; Gaps 14;

6 EPRVQAGLIGLGCGLGALVQLSLFSLAGLVAALVQVSPSSLSQEGSDAIVQNL 65

47 QPLAQRICSNVCF-----LLATSFNLL--LVVICV----- 76

66 TOLKAVGELSEKSLQEIYELTOLKAVGELPEKSLQEIYELRLKAVGELPEKS 125

77 -----TGSSEGHGAGLOALDLKSLKAFSNF--SSSTLTFEV-QAISTHGSGVCD----- 123

124 KLOEIYELRLKAVGELPEKSLQEIYELT-----FLAAVGEPLDQSKQOQIY 177

124 -----KITSLGA----- KLEKQOQDLKADHDLFLHMHFVULKRFVACQMLL 167

178 UHLTOLKTAERLCRCRCRQKMTFFQNCYFMSNSQRMWHDVTAQCEVPAQIVYIKTAE 237

168 H-----SNGSQRTG--CPVWVEHGSGYMFSGKAMAEAKYQGLENAHLYVINSNEE 220

238 UNFLOLQTSNPNPMSWGLSTINQETIMQWVDSPLSPSPQPVYNNSPFPNN-----SGN 231

221 UKLIVQHINFPN--TWGLTID--SDGSMKWVDPTDYPHNKN--WAVTQPDWGHGHELOGS 275

292 EDCAEFGSG--WMDNRCDVNNYICKK 317

276 EDCVEVQPDGRWMDDECTLYYRWYCK 302

RESULT 13

asialoglycoprotein receptor - mouse

N:Alternate names: hepatic lectin

C:Species: Mus musculus (house mouse)

C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Aug-1999

C:Accession: S13165

R:Sanford, J.P.; Doyle, D

Biochim. Biophys. Acta 1087, 259-261, 1990

A:Title: Mouse asialoglycoprotein receptor cDNA sequence, conservation of receptor genes

A:Reference number S13165, MIM:91027942; PMID:2233888

A:Molecule type: mRNA

A:Residues: 1-301 <SAN>

A:Cross-references EMBL X7042, NID:q53104, FIDR:MA172111, PTD:q53105

C:Superfamily: hepatic lectin; C-type lectin homology

C:Keywords: glycoprotein; liver; transmembrane protein

F:170-293/Domain: C-type lectin homology <LCH>

Query Match 15.24; Score 258; DB 2; Length 301;

Best Local Similarity 34.14; Pred. No. 86-093;

Matches 59; Conservative 28; Mismatches 56; Indels 30; Gaps 7;

170 QSKQQLYQHLT-----DLKTAERLC-----RHCRKMTFFQNCYFMSNS 211

128 EKKQQLKADHDLFLHMHFVULKRFVACQMLL----- 76

212 QRMWHDVTAQCEVPAQIVYIKTAEQNLQSLSPSPQPVYNNSPFPNN-----SGN 231

188 GLTWADQYQGLENAHLYVINSNEEDEF--VYKHSQFHITGLTID--KUGSMKWVDGT 243

272 PLSPSPQPVYNNSPFPNN-----SGNEDCAEFGSG--WMDNRCDVNNYICKK 317

244 DYKSNY--RMWAFQPDGRWMDDECTLYYRWYCK 295

RESULT 14

INTR2

hepatic lectin 2 - rat

N:Alternate names: asialoglycoprotein receptor PHL-2/3 (ASGP-2/3)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1988 #sequence_revision 09-Apr-1998 #text_change 22-

C:Accession: E64622, A64622, A16011, A26588, A25417

R:Halberg, D.F.; Wager, R.E.; Farrell, D.C.; Hildreth IV, J.; Jones

J. Biol. Chem. 263, 9828-9838, 1987

A:Title: Major and minor forms of the rat liver asialoglycoprotein

A:Reference number: A25417; MIM:872505b; PMID:6597443

A:Molecule type: mRNA

A:Residues: 1-301 -HAT-

A:Cross-references: GR:102762; NID:3205152, FIDR:MAA152211; PTD:q42

A:Accession: A28462

A:Molecule type: protein

A:Residues: 88-95, 'X', 98-118, 'X', 120, 129-159, 177-182, 'X', 184, 'X', 186

R:Sanford, J.P.; Elliott, R.W.; Doyle, D.

DNA 7, 721-728, 1988

A:Title: Asialoglycoprotein receptor genes are linked on chromosome

A:Reference number: A1601; MIM:8817011; PMID:4234178

A:Molecule type: mRNA

A:Residues: 1-301 -SAN-

A:Cross-references: GR:X6766, NID:q56000, FIDR:MAA504761; PTD:q42

R:McMullin, M.; Berg, P.

Mol. Cell. Biol. 7, 1841-1847, 1987

A:Title: Identification and characterization of cDNA clones encoding

A:Reference number: A25417; MIM:8725985; PMID:3600647

A:Molecule type: mRNA

A:Residues: 1-152, 'A', 154-201, 'T', 203-259, 'C', 261-301 -MCP-

A:Cross-references: GR:M16347; NID:q20646; FIDR:AAA420381; PTD:q42

A:Note: The authors translated the codon GCA for residue 153 as Arg

R:Dieckmann, K.; Mamont, J.F.; Bins, G.; Leung, J.O.

J. Biol. Chem. 259, 770-778, 1984

A:Title: Primary structure of the rat liver asialoglycoprotein recep

A:Reference number: A25417, MIM:84111554, PMID:6319386

A:Molecule type: protein

A:Residues: 201-259, 'C', 261-281, 'ND', 281-301 -DRI-

C:Comment: Calcium is required for ligand binding.

C:Superfamily: hepatic lectin; C-type lectin homology

C:Keywords: endocytosis; glycoprotein; lectin; receptor; LRR

F:2-60/Domain: intracellular #status predicted -INT2

F:61-77/Domain: transmembrane #status predicted -EXT2

F:78-301/Domain: extracellular #status predicted -EXT2

F:170-293/Domain: C-type lectin homology <LCH>

F:97,119,165/Binding site: carbohydrate (Asu) (covalent) #status p

Query Match 14.38; Score 243; DB 1; Length 301;

Best Local Similarity 37.44; Pred. No. 6,58-08;

Matches 49; Conservative 22; Mismatches 48; Indels 12

194 CPKMTFFQNCYFMSNSQRMWHDVTAQCEVPAQIVYIKTAEQNLQSLSPSPQPVYNNSPFPNN-----SGNEDCAEFGSG--W 211

170 CPVWVEHGSGYMFSGKAMAEAKYQGLENAHLYVINSNEEDEF--VYKHSQFHITGLTID--KUGSMKWVDGT 243

254 MGLSDLNQEGTQWVDSPLSPSPQPVYNNSPFPNN-----SGNEDCAEFGSG--W 211

228 LGLLD--KUGSMKWVDGTFTYNSNFKN--WAFQPDGRWMDDECTLYYRWYCK 295

307 CDVDNYICKK 317

285 CQGVNRYACKER 295

RESULT 15

UC7608

type II lectin-like immunoreceptor - human

C:Species: Homo sapiens (man)

C:Date: 30-Jun-2001 #sequence_revision 10-Jun-2001 #text_change 10-

C; Accession: JC7608

K; Huang, X., Yuan, Z., Chen, G., Zhang, M., Zhang, W., Yu, Y., Cao, X., Zhou, B. 2001. 131-140

Biochem. Biophys. Res. Commun. 281, 131-140, 2001

A: Title: Cloning and characterization of a novel TIM containing lectin-like immunorecep

A: Contents: dendritic cells
A/relief: none number: 0070000, model: 21002/977, PMID: 1118971

A; Accession: JC7608

A; molecule type: mRNA

A; residues: 1-237 <HUA>

A: Cross-references: GB:

Comment: This receptor, highly conserved in invertebrates

cell, especially in migrating, antigen capturing and processing.

A:Gene: Mir

A: Map position

11:45-69/Domain: Trans

Query Match:	13.78;	Score	232;	DB 2;
Host Local Similarity:	34.98;	Prod No	3	30-07-

Best Local Similarity	34.88;	Pred. No: 2.3e-07;
Matches	55	Conserved Iw: 30
Mismatches	53	Totals

Match	Conservative	Mismatches
1	20	33
2	20	33
3	20	33
4	20	33
5	20	33
6	20	33
7	20	33
8	20	33
9	20	33
10	20	33
11	20	33
12	20	33
13	20	33
14	20	33
15	20	33
16	20	33
17	20	33
18	20	33
19	20	33
20	20	33
21	20	33
22	20	33
23	20	33
24	20	33
25	20	33
26	20	33
27	20	33
28	20	33
29	20	33
30	20	33
31	20	33
32	20	33
33	20	33
34	20	33
35	20	33
36	20	33
37	20	33
38	20	33
39	20	33
40	20	33
41	20	33
42	20	33
43	20	33
44	20	33
45	20	33
46	20	33
47	20	33
48	20	33
49	20	33
50	20	33
51	20	33
52	20	33
53	20	33
54	20	33
55	20	33
56	20	33
57	20	33
58	20	33
59	20	33
60	20	33
61	20	33
62	20	33
63	20	33
64	20	33
65	20	33
66	20	33
67	20	33
68	20	33
69	20	33
70	20	33
71	20	33
72	20	33
73	20	33
74	20	33
75	20	33
76	20	33
77	20	33
78	20	33
79	20	33
80	20	33
81	20	33
82	20	33
83	20	33
84	20	33
85	20	33
86	20	33
87	20	33
88	20	33
89	20	33
90	20	33
91	20	33
92	20	33
93	20	33
94	20	33
95	20	33
96	20	33
97	20	33
98	20	33
99	20	33
100	20	33

175 Q1YQETD1KTAFFR1-----CRHCPKDWTFEÖGNCYFMSNSÖRNM 215
ÜY

[illegible]

70 QKXSÜLLEKKTKEI.VHTT.ECVKKNMVEETAMSC - -OPKNWKSFSŠNČYFIŠTESASW 127

316 UNIVERSITY OF WYOMING DEPARTMENT OF CHEMISTRY, BOX 3336, LARAMIE, WYOMING 82071-3336

216 HDSTACQFVKAQI.VIKTAEQNF1.QLOTSKSNKFSWGLSLDLNQEGTWQWVDGSP 275

128 OJSEKIC ARMEALVIN OF EOPH 1-FON OESAY EVIS

Q. Now, you're going to tell me that the defendant was not in the car at the time of the shooting, is that correct?

276. $\text{SFOPYWNSGHINNSGNEGTA} \cdots \text{EFGS} \cdots \text{GWNINPC}$ 307

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 11:27:01 : Search time 2718 Seconds
(without alignments)
3479,916 Million cell updates/sec

Title: US-09-831-458A-12
Perfect score: 1696
Sequence: 1 MDSKEPPVQQLSLGCTICH : KCIWDNWIQCKPACFRFR 325

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame_plus_model -PEV-yih
-Q/can2_1/USPTO_SPOOL/050861458/umat.0512000:10319_15247/46f_query.fasta.1.515
-DB-GenEmbl -OFMT=fastap -SUFFIX=std.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cbl -LIST=45
-DOCALLON=200 -THR_SCORE=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OFFMT=prt -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=us09831458.us09.1.1010.stdat_0512000_10319_15247 -RCPPI=6 -ICPI=3
-NO_XLPRY -NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -LON3146 -PEV_TIMECUT=120
-WARN_TIMECUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: GenEmbl:
1: gb_fa:
2: gb_hg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pa:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_tv:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vl:
15: em_fa:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pa:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:

29: em_vl:
30: em_hlg_hum:
31: em_hlg_inv:
32: em_hlg_other:
33: em_hlg_mus:
34: em_hlg_pin:
35: em_hlg_rtd:
36: em_hlg_mam:
37: em_hlg_vfl:
38: em_sy:
39: em_hlgo_hum:
40: em_hlgo_mus:
41: em_hlgo_other:

Pred. No. is the number of results predicted by chance to be
score greater than or equal to the score of the result being
and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Query Match	Length	DB	ID	Descr
1	1646.5	97.1	1800	9	AF290887	AF290887
2	1639	96.6	1200	9	AY042234	AY042234
3	1628	96.0	1510	9	AF245219	AF245219
4	1536.5	90.6	1648	9	AY042235	AY042235
5	1494.5	88.1	1212	6	AR166105	AR166105
6	1412.5	83.3	1212	6	AX342003	AX342003
7	1412.5	83.3	1215	6	AX039403	AX039403
8	1412.5	83.3	1215	6	AY042221	AY042221
9	1412.5	83.3	1312	6	AX287103	AX287103
10	1412.5	83.3	1312	6	HDMLECTINC	HDMLECTINC
11	1412.5	83.3	1643	6	AF290886	AF290886
12	1412.5	83.3	1643	6	AX287118	AX287118
13	1407	83.0	1643	9	AB015629	AB015629
14	1407	82.8	1212	9	AF361085	AF361085
15	1404.5	82.5	1212	6	AF166106	AF166106
16	1398.5	82.5	1212	6	AX342004	AX342004
17	1398.5	82.5	1212	6	AF166105	AF166105
18	1394.5	82.4	1212	6	AX342005	AX342005
19	1393.5	82.2	1212	6	AX342006	AX342006
20	1378	81.2	1315	9	AY042229	AY042229
21	1375	81.1	1146	9	AF369755	AF369755
22	1375	81.1	1166	9	AY040319	AY040319
23	1369.5	80.7	1197	9	AY042222	AY042222
24	1365	80.5	1146	9	AF343727	AF343727
25	1347.5	79.5	1215	9	AF391086	AF391086
26	1342	79.1	1105	9	AY042231	AY042231
27	1315.5	77.6	1083	9	AY042226	AY042226
28	1314.5	77.5	1143	9	AY042225	AY042225
29	1313	77.4	1243	9	AY042230	AY042230
30	1280.5	75.5	1535	9	AY042236	AY042236
31	1215.5	71.7	939	9	AY042237	AY042237
32	1151	67.9	811	9	AY042238	AY042238
33	1148	67.7	792	9	AY042239	AY042239
34	1144	67.5	1472	9	AY042239	AY042239
35	1112.5	65.6	807	9	AC008812	AC008812
36	1078.5	63.6	143619	9	AF209479	AF209479
37	1065.5	62.8	4506	2	AF209480	AF209480
38	1054.5	62.2	190628	2	AF209480	AF209480
39	971.5	57.3	2991	9	AY042232	AY042232
40	964.5	56.9	994	9	AF209479	AF209479
41	883.5	52.1	4424	9	AC008812	AC008812
42	877	51.7	143619	9	AC008812	AC008812
43	877	51.7	177962	9	AF373409	AF373409
44	720	43.5	1220	10	AF373409	AF373409
45	631	37.2	927	10	AF373471	AF373471

RESULT 1

ALIGNMENTS


```

FEATURES
Source      Location/Qualifiers
1         1212
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/cell_type="monocyte-derived dendritic cells"
1..>1212
/gene="DC-SIGN"
1..>1212
/gene="DC-SIGN"
/note="HIV-1 trans-receptor, CD209"
/codon_start=1
/product="dendritic cell-specific ICAM-3 grabbing
non-integrin"
/protein_id="AAK97458.1"
/db_xref="GI:15420782"
/translation="MSDSKEPRLOQIGLEBEULKQGFQNGYKSLAGIGHPLV
LQLSFTLLAGLVKSPSSISOESRODYIYONITOLKAAGELSEKSLQELIYQ
ELTOLKAAGELPEKSKQELIYQELTRKAAGELPEKSKQELIYQELTRKAAGEL
PEKSKQELIYQELTRKAAGELPEKSKQELIYQELTRKAAGELPEKSKQELIYQ
LTRKAAGELPEKSKQELIYQELTOLKAAGELPEKSKQELIYQELTOLKAAGEL
WDSITACKRVGQIVIKSAEONFLOQSSRSNFTWMLSDINEGTQWQWDGSP
LIPSFQYNNKCEPNNVCEDECAEISGNMNDKCNLAKFWLCKNSAASCSHDEDFL
SPAPATPNPPPA"
BASE COUNT      315 a      316 c      351 g      230 t
ORIGIN
Alignment Scores:
Pred No.      8 72e-102      Length:      1212
Score:      1404.50      Matches:      278
Percent Similarity:      76.79%      Conservative:      23
Best local Similarity:      70.92%      Mismatches:      20
Query Match:      82.81%      Indels:      71
Db:      5      Gaps:      4
US-09-831-458a-12 (1-325) x AF391085 (1-1212)
QY      1 MetSerAspSerLysGluProArgValGlnGlnLeuGlyLeuLeu----- 15
Db      1 ATGAGTGTACTCCAGCAAGCAAGCTGCAGCAGCTGGGCTTCGAGGAGGACAGCTG 60
QY      16 -----GlycylsLeuGlyHis 20
Db      61 AGAGGCTTGGATTCCGACACAAATCGAGGCTTACAAAGCTTACGAGGCTTCGGCCAT 120
QY      21 GYAlaLeuValLeuGlnLeuLeuSerPheMetLeuLeuAlaGlyValLeuValAlaIle 40
Db      121 GCTCCCTCGTGGTGGTCAACACGCTCTCTTACGCTCTTGGCTGGG-----CTC 168
QY      41 LeuValGlnValSerLysValProSerSerLeuSerGlnGlnGlnSerGlnGlnAspAla 60
Db      169 CTTCCTCAAGCTGCAAGGCTTCTTAACTTCAATAATAGTAAAGAAATCAAGACGAG 228
QY      61 LLeTyrGlnAsnLeuThrGlnLeuLysAlaAlaValGlyGlnLeuSerGlnLysSerLys 80
Db      229 ATTACATCAAACTGACCTTAAAGTTGTAATGAGTAACTTCAAGAAATCAAG 288
QY      81 LeuGlnGlnLeuTyrGlnGlnLeuThrGlnLeuLysAlaAlaValGlyGlnLeuProGln 100
Db      289 CTGCAGAGATCTACAGCAAGCTGACCTCAAGGCTGCTGAGCTTCCAGAG 348
QY      101 LysSerLysLeuGlnGlnLeuTyrGlnGlnLeuThrArgLeuLysAlaAlaValGlyGln 120
Db      349 AAATCTAAGCAGCAGGATCTACAGGAGCTGACCCGCTCAAGGCTGCTGAG 408
QY      121 LeuProGlnLysSerLysLeuGlnGlnLeuTyrGlnGlnLeuThrArgLeuLysAlaAla 140
Db      409 CTTCAGAGAAATCTAGATGAGAGATCTTACCTAGGAGTGTGACTGGGCTGAAGGCTGCA 468
QY      141 ValGlyGlnLeuProGlnLysSerLysLeu----- 150
Db      469 GTGGGTGAGCTTCCAGAGAAATCTAAGATCTAGTAACTTACAGAGAAATCTGACTGGCTG 528
QY      150 ----- 150

```

```

Db      529 AACCTTCAATGCTGATGATCTTCTGCAAAATCAAGCAAGGAGATCTG 100
QY      151 ----- 151
Db      589 ACCGAGTTGAAGGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 100
QY      155 GlnGlnLeuThrArgLeuLysAlaAlaValGlyGlnLeuProGln 100
Db      649 CAAAGGTTGACCGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 100
QY      175 GlnLysGlnGlnLeuThrAspLeuLysThrAlaPheGlnArgLeu 100
Db      709 GAGATCTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 100
QY      195 ProLysAspTrpThrPhePheGlnGlnLysAsnGlySerPheMetSerAsn 100
Db      769 CCTGGGATGACATCTCTTCAAGGAAATGTTACTTCAATGCTTAACT 100
QY      215 TrpHisAspSerValThrAlaGlyGlnGlnValArgAlaGlnLeuVal 100
Db      829 TGGCAGTAACTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 100
QY      235 AlaGlnGlnGlnAsnPheLeuGlnGlnLeuThrSerArgSerAsnArg 100
Db      889 GTTAAAGAAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 100
QY      255 GlyLeuSerAspLeuAsnGlnGlnGlyThrTrpSerTrpValAspGlySer 100
Db      949 GCACTTCAATCTAAATGAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 100
QY      275 ProSerPheGlnArgTyrTrpAsnSerGlyGlnProAsnAsnSerGlyVal 100
Db      1009 CCAAGCTTCAATCTAAATGAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 100
QY      295 AlaGlnPheSerGlySerGlyTrpAsnAspAsnArgGlyAspValAspAla 100
Db      1069 GCGAATTTAGTGGCAATGCTGCAATGCTGCAATGCTGCAATGCTGCAATGCTG 100
QY      315 CysLysLysProAlaAla-----CysPheArgAspGln 325
Db      1129 TGCAAAAGTCCGAGGCTCTCTCTTCAAGGATGAA 1164

```

Search completed: December 7, 2002, 14:12:41
Job time : 2737 secs

GenCore version 5.1.3
Copyright (c) 1993 2002 Compaq Inc

OM protein - nucleic search, using frame_plus_p2n model

Run on December 7, 2002, 11:28:11, Search time 2060 Seconds

(without alignments)
2555.113 Million cell updates/sec

Title: US-09-831-458a-12
Percent score: 1696
Sequence: 1 MSDSKPRVQUGLGLGICG.....RCIVDWYICKRACERDE 325

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32398132

Minimum Dh seq length: 0
Maximum Dh seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n model; DEY=xlh
-O=/cgn2-1/OSPro/spool/US09831458/runat_05122002_103219_17283/ffg_jetty.fasta_1.519
-DB=EST-qfwt-fastlap -SUFFIX=std_rst -MINMATCH=0.1 -DOOPCT=0 -DOOPEXT=0
-UNITS=bits -STAR1=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=PL0 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=0509831458.#CGN.1.1.763 -runat_05122002_103219_17283 -NCPU=6 -TFTP=3
-NO_XLPYX -NO_MMAP -LARGEMEMORY -NCR=SCORES=0 -WAIT -CONVLOG -DEY_TIMEOUT=120
-WARN_TIMEOUT=10 -THRAPAS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -Delop=6 -DELEXT=7

Database: EST:
1: em_estba:
2: em_esthum:
3: em_estim:
4: em_estim:
5: em_estim:
6: em_estim:
7: em_estim:
8: em_estim:
9: qd_est1:
10: qd_est2:
11: qd_est3:
12: qd_est4:
13: qd_est5:
14: qd_est6:
15: em_estfun:
16: em_estim:
17: qd_gss:
18: em_gss_hum:
19: em_gss_hum:
20: em_gss_hum:
21: em_gss_hum:
22: em_gss_hum:
23: em_gss_hum:
24: em_gss_hum:
25: em_gss_hum:
26: em_gss_hum:
27: em_gss_hum:

Pred. No. is the number of results predicted by chance
score greater than or equal to the score of the result
and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	Dh	Id	Id
1	849	50.1	792	9	AL544658	AL544658
2	827.5	48.8	828	10	BE562964	BE562964
3	827	48.8	769	13	B1828875	B1828875
4	763	45.0	786	13	B1827056	B1827056
5	739	43.6	706	10	AV445756	AV445756
6	714	42.1	503	14	R98113	R98113
7	711.5	42.0	1136	11	AK007654	AK007654
8	710	41.9	792	9	AI140165	AI140165
9	623	39.7	669	10	BM369427	BM369427
10	559.5	35.0	654	13	BM369428	BM369428
11	520.5	30.7	441	9	AA914211	AA914211
12	434.5	25.6	655	10	BM683702	BM683702
13	428.5	25.3	1165	11	AK007638	AK007638
14	424.5	25.0	510	13	BM462374	BM462374
15	424	25.0	833	9	AL570816	AL570816
16	421.5	24.9	438	9	AL507520	AL507520
17	417	24.6	1045	11	AK010071	AK010071
18	407.5	24.0	614	10	BM637496	BM637496
19	400	23.6	419	14	BE90968	BE90968
20	394	23.2	490	9	AL548311	AL548311
21	391	23.1	359	19	AV445629	AV445629
22	384	22.6	1211	13	BE921322	BE921322
23	379.5	22.4	398	9	AA510960	AA510960
24	369.5	21.8	284	9	AL614060	AL614060
25	358.5	21.1	718	10	BE374084	BE374084
26	343	20.2	566	9	AA543874	AA543874
27	330.5	19.5	633	10	BM665930	BM665930
28	330	19.5	395	9	AL438251	AL438251
29	323	19.0	486	14	BM014470	BM014470
30	308	18.2	1197	11	BM014479	BM014479
31	302.5	17.8	429	9	AA165054	AA165054
32	296	17.5	434	9	AL286429	AL286429
33	292.5	17.2	413	9	BE363665	BE363665
34	291.5	17.0	557	12	BM712387	BM712387
35	289	17.0	816	14	BM018186	BM018186
36	287	16.9	953	14	BE146245	BE146245
37	277.5	16.4	597	12	BE532369	BE532369
38	275	16.2	774	13	BE182064	BE182064
39	274	16.2	500	9	AA048806	AA048806
40	273.5	16.1	738	14	BM004812	BM004812
41	272.5	16.1	569	12	BE182065	BE182065
42	272.5	16.1	945	14	BM446882	BM446882
43	268.5	15.8	574	12	BM410604	BM410604
44	268.5	15.8	736	13	BM104514	BM104514
45	258	15.8	816	12	BE755884	BE755884

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AL544658	AL544658	AL544658 LTI-NFL006.F12 Homo sapiens cDNA clone	AL544658	AL544658	EST	human	Homo sapiens	Li, W.B., Gruber, C., Jessen, J. and Polayes, D.	Full-length cDNA libraries and normalization	
AL544658	AL544658	AL544658 LTI-NFL006.F12 Homo sapiens cDNA clone	AL544658	AL544658	EST	human	Homo sapiens	Li, W.B., Gruber, C., Jessen, J. and Polayes, D.	Full-length cDNA libraries and normalization	

[illegible]

RESULT 9	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
BB636927	BB636927	BB636927 FIKEN full-length cDNA, adult male testis and ovary muscle cDNA clone A510045B4.5, mRNA sequence.	BB636927	BB636927	BB636927.1	GI:1472713	house mouse, <i>Mus musculus</i>

REFERENCE
1. (Pages 1 to 669)
Atakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, T., Komori, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, K., Sakai, Y., Sakai, Y., Sasaki, Y., Sakurai, M., Shihara, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, T., Takeya, Y., Tanabe, T., Toyota, T., Mitamura, M., and Hayashizaki, T. (2001)
P1KEN MOUSE ESTS (Atakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshihisa Hayashizaki

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suchoiro-cho, Tsukuba, Ibaraki 305-0857, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@psc.riken.go.jp,
URL: <http://genome.psc.riken.go.jp/>
Garnick, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
M., Kono, H., Okazaki, Y., Moramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapped selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. *Genome Res.* 10 (10), 1617-1630 (2000)
Wagai, K., Fujiiwara, S., Imoto, K., Toyama, Y., Izawa, M., Chira, F.,
Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Imoto, Y., Kita, A.
and Hayashizaki, Y.

Kondo S., Shindagawa A., Saito T., Kiyosawa H., Yamamoto I., Akizawa K., Fukuda S., Hara A., Itoh M., Kawai J., Shibata K. and Hayashizaki Y. (2000) A new integrated sequence analysis (ISA) system for sequencing pipeline with 364 multicapillary sequencer. *Genom. Res.* 10 (11), 1757-1771.

Kondo S., Hayashizaki Y., Shibata K., Itoh M., Carninci P., Sugahara Y. and Hayashizaki Y. (2001) Computer based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289.

availability. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome*, 12, 673-677 (2001)

Please visit our web site (<http://genome.gsfc.nasa.gov>) for further details.

e mouse tissues.

location/qualifiers

1, 669

source

BASE COUNT	202 a	148 c	165 g	152 t	2 others
ORIGIN					
Alignment Scores:					
Fred. No.:	3,096-67				
Score:	624.90				669
Percent Similarity:	50.78%				126
Post Local Similarity:	39.50%				36
Query Match:	36,734				41
DB:	10				116
Caps:					4

[illegible]

[illegible]

AUTHORS	TITLE	JOURNAL	DOI
Grumkin, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibata, K., Itoh, M., Komori, H., Okazaki, Y., Watanabe, M. and Hayashizaki, Y.	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full length cDNA libraries for rapid discovery of new genes	Genome Res. 10 (40), 1617-1630 (2000)	
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carolinci, P., Komori, H., Akiyama, J., Nishi, K., Kitasumi, T., Yoshino, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, K., Yamashita, R., Matsushima, H., Sakayuchi, S., Ikegami, T., Kasai, Y., Fujikawa, S., Inoue, K., Iwagawa, Y., Izawa, M., Ohara, E., Matsuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer	Genome Res. 10 (11), 1757-1771 (2000)	
Kawai, J., Shitagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Harada, A., Fukunishi, Y., Komori, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamazaki, J., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., Kiro, H., Koehliwa, H., K. L. F., Lewis, S., Matsuo, Y., Nakajima, I., Perle, G., Quackenbush, J., Schriml, J. M., Stuppi, F., Suzuki, K., Tomita, M., Wagner, J., Washio, T., Sakai, K., Okido, T., Pirro, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Bottelli, D., Bojunga, N., Carolinci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Farnham, P., Fujita, M., Garibay, M., Gustinovich, S., Hill, P., Hironaka, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombardi, P., Nordone, P., Rhee, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schomburg, C., Segal, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, H., Wong, G., Whitaker, C., Wilting, J., Wynshaw-Boris, A., Yoshida, K., Basegawa, Y., Kawai, H., Kohnsaki, S. and Hayashizaki, Y.	Functional annotation of a full-length mouse cDNA collection	Nature 409 (6821), 685-690 (2001)	
Grumkin, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibata, K., Itoh, M., Komori, H., Okazaki, Y., Watanabe, M. and Hayashizaki, Y.	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full length cDNA libraries for rapid discovery of new genes	Genome Res. 10 (40), 1617-1630 (2000)	
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carolinci, P., Komori, H., Akiyama, J., Nishi, K., Kitasumi, T., Yoshino, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, K., Yamashita, R., Matsushima, H., Sakayuchi, S., Ikegami, T., Kasai, Y., Fujikawa, S., Inoue, K., Iwagawa, Y., Izawa, M., Ohara, E., Matsuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer	Genome Res. 10 (11), 1757-1771 (2000)	
Kawai, J., Shitagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Harada, A., Fukunishi, Y., Komori, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamazaki, J., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., Kiro, H., Koehliwa, H., K. L. F., Lewis, S., Matsuo, Y., Nakajima, I., Perle, G., Quackenbush, J., Schriml, J. M., Stuppi, F., Suzuki, K., Tomita, M., Wagner, J., Washio, T., Sakai, K., Okido, T., Pirro, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Bottelli, D., Bojunga, N., Carolinci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Farnham, P., Fujita, M., Garibay, M., Gustinovich, S., Hill, P., Hironaka, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombardi, P., Nordone, P., Rhee, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schomburg, C., Segal, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, H., Wong, G., Whitaker, C., Wilting, J., Wynshaw-Boris, A., Yoshida, K., Basegawa, Y., Kawai, H., Kohnsaki, S. and Hayashizaki, Y.	Functional annotation of a full-length mouse cDNA collection	Nature 409 (6821), 685-690 (2001)	
Grumkin, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibata, K., Itoh, M., Komori, H., Okazaki, Y., Watanabe, M. and Hayashizaki, Y.	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full length cDNA libraries for rapid discovery of new genes	Genome Res. 10 (40), 1617-1630 (2000)	
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carolinci, P., Komori, H., Akiyama, J., Nishi, K., Kitasumi, T., Yoshino, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, K., Yamashita, R., Matsushima, H., Sakayuchi, S., Ikegami, T., Kasai, Y., Fujikawa, S., Inoue, K., Iwagawa, Y., Izawa, M., Ohara, E., Matsuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer	Genome Res. 10 (11), 1757-1771 (2000)	
Kawai, J., Shitagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Harada, A., Fukunishi, Y., Komori, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamazaki, J., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., Kiro, H., Koehliwa, H., K. L. F., Lewis, S., Matsuo, Y., Nakajima, I., Perle, G., Quackenbush, J., Schriml, J. M., Stuppi, F., Suzuki, K., Tomita, M., Wagner, J., Washio, T., Sakai, K., Okido, T., Pirro, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Bottelli, D., Bojunga, N., Carolinci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Farnham, P., Fujita, M., Garibay, M., Gustinovich, S., Hill, P., Hironaka, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombardi, P., Nordone, P., Rhee, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schomburg, C., Segal, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, H., Wong, G., Whitaker, C., Wilting, J., Wynshaw-Boris, A., Yoshida, K., Basegawa, Y., Kawai, H., Kohnsaki, S. and Hayashizaki, Y.	Functional annotation of a full-length mouse cDNA collection	Nature 409 (6821), 685-690 (2001)	
Grumkin, P., Shibata, Y., Hayatsu, N., Sugihara, Y., Shibata, K., Itoh, M., Komori, H., Okazaki, Y., Watanabe, M. and Hayashizaki, Y.	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full length cDNA libraries for rapid discovery of new genes	Genome Res. 10 (40), 1617-1630 (2000)	
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carolinci, P., Komori, H., Akiyama, J., Nishi, K., Kitasumi, T., Yoshino, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, K., Yamashita, R., Matsushima, H., Sakayuchi, S., Ikegami, T., Kasai, Y., Fujikawa, S., Inoue, K., Iwagawa, Y., Izawa, M., Ohara, E., Matsuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer	Genome Res. 10 (11), 1757-1771 (2000)	
Kawai, J., Shitagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Harada, A., Fukunishi, Y., Komori, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamazaki, J., Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., Kiro, H., Koehliwa, H., K. L. F., Lewis, S., Matsuo, Y., Nakajima, I., Perle, G., Quackenbush, J., Schriml, J. M., Stuppi, F., Suzuki, K., Tomita, M., Wagner, J., Washio, T., Sakai, K., Okido, T., Pirro, M., Aono, H., Balderelli, R., Barsh, G., Blake, J., Bottelli, D., Bojunga, N., Carolinci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Farnham, P., Fujita, M., Garibay, M., Gustinovich, S., Hill, P., Hironaka, M., Hume, D. A., Kamuya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombardi, P., Nordone, P., Rhee, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schomburg, C., Segal, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, H., Wong, G., Whitaker, C., Wilting, J., Wynshaw-Boris, A., Yoshida, K., Basegawa, Y., Kawai, H., Kohnsaki, S. and Hayashizaki, Y.	Functional annotation of a full-length mouse cDNA		

is vector free and a least 200bp in length
Insert length: 510 Std Error: 0.00
Plate: BS220044060 row: B column: 5
Seq primer: CGCCAACTCTCAAAATTAACCC
High quality sequence start: 510.
Location/Qualifiers

BASE COUNT	126 a	127 c	145 q	112 t
ORIGIN				

US-09-831-458A-12 (1-325) x BM65374 (1-510)	
Presequence Gaps:	
Pred. No.:	6,93a-40
Score:	424.50
Percent Similarity:	63.41%
Best Local Similarity:	45.71%
Query Match:	25.03%
DB:	13
	Gaps:
	2
	Length:
	510
	Matches:
	75
	Overlapped:
	25
	Mismatches:
	55
	Indels:
	5

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 11:17:46 - Search time 277 seconds

(without alignments)
2642.236 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 1696
Sequence: 1 MSISKEPRVQGLILGCTGH PCHVNWYICPKPACPFIDE 325

Scoring table:

BLUSTR62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delpept 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2-1/ospo/spool/US081458/rnmat_05122002_103215_15129/ff-1query.fasta.1.519
-DB=N_Geneseq_101002 -OPMT=fastap -SUFFIX=std.rng -MINMATCH=0.1 -L=0pct.0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blustr62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTEXT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09831458.rcgn.1.1.0.rnmat_05122002_103215_15129 -NFW=6 -ICPU=3
-NO_XLPHY -NO_MMAR -LARGEJURY -NRG_SCORES=6 -WAIT -L7NFWING -FEV_TIMEOUT=120
-WARN_TIMEOUT=10 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

1. N_Geneseq_101002.*
2. /SID52/gcdata/geneseq/geneseqn-emb1/NA1980.LAT.*
3. /SID52/gcdata/geneseq/geneseqn-emb1/NA1981.LAT.*
4. /SID52/gcdata/geneseq/geneseqn-emb1/NA1982.LAT.*
5. /SID52/gcdata/geneseq/geneseqn-emb1/NA1983.LAT.*
6. /SID52/gcdata/geneseq/geneseqn-emb1/NA1984.LAT.*
7. /SID52/gcdata/geneseq/geneseqn-emb1/NA1985.LAT.*
8. /SID52/gcdata/geneseq/geneseqn-emb1/NA1986.LAT.*
9. /SID52/gcdata/geneseq/geneseqn-emb1/NA1987.LAT.*
10. /SID52/gcdata/geneseq/geneseqn-emb1/NA1988.LAT.*
11. /SID52/gcdata/geneseq/geneseqn-emb1/NA1989.LAT.*
12. /SID52/gcdata/geneseq/geneseqn-emb1/NA1990.LAT.*
13. /SID52/gcdata/geneseq/geneseqn-emb1/NA1991.LAT.*
14. /SID52/gcdata/geneseq/geneseqn-emb1/NA1992.LAT.*
15. /SID52/gcdata/geneseq/geneseqn-emb1/NA1993.LAT.*
16. /SID52/gcdata/geneseq/geneseqn-emb1/NA1994.LAT.*
17. /SID52/gcdata/geneseq/geneseqn-emb1/NA1995.LAT.*
18. /SID52/gcdata/geneseq/geneseqn-emb1/NA1996.LAT.*
19. /SID52/gcdata/geneseq/geneseqn-emb1/NA1997.LAT.*
20. /SID52/gcdata/geneseq/geneseqn-emb1/NA1998.LAT.*
21. /SID52/gcdata/geneseq/geneseqn-emb1/NA1999.LAT.*
22. /SID52/gcdata/geneseq/geneseqn-emb1/NA2000.LAT.*
23. /SID52/gcdata/geneseq/geneseqn-emb1/NA2001.LAT.*
24. /SID52/gcdata/geneseq/geneseqn-emb1/NA2002.LAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID
1	1696	100.0	1542	AAA27055
2	1412.5	83.3	1215	AAV41224
3	1412.5	83.3	1215	AAA88740
4	1412.5	83.3	1215	AAV75484
5	1412.5	83.3	1215	AAV65461
6	1407	83.0	1643	AAV84361
7	1407	83.0	1643	AAV65469
8	1398.5	82.5	1212	AAV41224
9	1398.5	82.3	1312	AAV66448
10	1393.5	82.3	1212	AAV41225
11	1393.5	82.3	1212	AAV41225
12	971.5	57.3	2291	AAH16656
13	710	41.9	792	AAH08239
14	545.5	32.2	1516	AAH22221
15	527	31.1	433	AAH9365
16	515	30.4	553	AAH07709
17	415	24.5	909	AAH45062
18	350	20.6	1005	AAH41954
19	350	20.6	1053	AAH48023
20	350	20.6	1064	AAH50144
21	337	19.9	885	AAH41957
22	332.5	19.6	924	AAH43958
23	329	19.4	1531	AAH43957
24	325.5	19.2	1504	AAH41485
25	325.5	19.2	1531	AAH70107
26	322.5	19.0	1037	AAH41955
27	322.5	19.0	1494	AAH6344
28	322.5	19.0	1503	AAH21303
29	322.5	19.0	1503	AAH45181
30	322.5	19.0	1504	AAH81367
31	322.5	19.0	1515	AAH20134
32	322.5	19.0	1529	AAH81613
33	322.5	19.0	1530	AAH21304
34	322.5	19.0	1530	AAH45182
35	322.5	19.0	1532	AAH41305
36	322.5	19.0	1532	AAH45184
37	319.5	18.8	1025	AAH41956
38	319.5	18.8	1458	AAH4866
39	318.5	18.8	1504	AAH45112
40	317	18.7	1479	AAH4138
41	317	18.7	1549	AAH4252
42	305	18.0	1370	AAH4868
43	302.5	17.8	1465	AAH4137
44	296.5	17.5	1355	AAH43685
45	296.5	17.5	1355	AAH78506

ALIGNMENTS

RESULT 1	AAA27055	standard; CDNA: 1512 bp
ID	AAA27055	
AC	AAA27055	
XX		
XX		
XX		
XX		
DE	22-AUG-2000	(first entry)
XX		
XX		
KW	Human: HCSRPT: cytosolic; antiarthritic; antirheumatic;	
KW	immunoprotective; footrot; antitubercular; cancer; le-	
KW	metastoma; rheumatoid arthritis; asthma; atherosclerosis;	
KW	Alzheimer's diseases; multiple sclerosis; epilepsy; ss	
OS	Homo sapiens.	
XX		


```

XX 06-JUN-2001: 2001W0-0518310.
XX
XX
XX 09-JUN-2000: 2000W0-0591435.
XX
XX (EVOL-) EVOLUTIONARY GENOMICS LLC.
XX
XX Messier W, Sikela JM.
XX
XX WPI: 2002-130744/17.
XX
XX Determining human or non-human primate polynucleotide or polypeptide
XX sequences associated with a physiological trait and have undergone
XX evolutionary changes, for therapeutic use, involves using statistical
XX methods.
XX
XX Example 19; Fig 11; 14pp; English.
XX
XX The present sequence is that of the coding region of the human
XX DC-SIGN gene. DC-SIGN is expressed on dendritic cells and is known
XX to provide a mechanism for transport of HIV-1 virus to the lymph
XX nodes. HIV-1 binds to the extracellular portion of DC-SIGN and
XX infects the undifferentiated T cells in the lymph nodes via their
XX CD4 proteins. This ultimately leads to compromise of the immune
XX system and to full-blown AIDS. The invention comprises a
XX comparative genomics approach to identify specific gene changes
XX responsible for differences in functions and diseases distinguishing
XX humans from non-humans, particularly primates including gorilla,
XX orangutan and especially chimpanzee. The evolutionary significance
XX of a nucleotide change is determined by the ratio of the
XX non-synonymous substitution rate (KA) to the synonymous rate (KS)
XX of the nucleotide sequence. Polynucleotide and polypeptide
XX sequences corresponding to evolved traits may be relevant to human
XX diseases or conditions such as unique or enhanced human brain
XX functions, longer human life spans, susceptibility or resistance to
XX disease, including AIDS and cancer, and aesthetic traits such as
XX hair growth. Ka/Ks ratios for chimpanzee/human, human/gorilla and
XX chimpanzee/gorilla DC-SIGN cDNA sequences are 1.3, 0.87 and 1.3,
XX respectively. It is theorized that chimpanzee resistance to
XX progression to full-blown AIDS may be due in part to inability of
XX HIV-1 to bind to chimpanzee DC-SIGN for transport to the lymph
XX nodes. After determining the 3-dimensional structure of DC-SIGN,
XX a rational drug design approach can be used to mimic the effects
XX of chimpanzee DC-SIGN without interfering with the normal functions
XX of human DC-SIGN.
XX
XX Sequence 1212 BP: 314 A, 321 C, 348 G, 229 T, 0 other:
XX
XX Alignment Scores:
XX Pred No. 1 12a-115 Length: 1212
XX Score: 1412.50 Matches: 282
XX Percent Similarity: 77.04% Conservative: 20
XX Best Local Similarity: 71.94% Mismatches: 19
XX Query Match: 83 28% Inpct: 71
XX DB: 24 Gaps: 4
XX
XX US-09-831-458a-12 (1-325) x ABA91223 (1-1212)
XX
XX 1 MetSerAspSerLysGluProAlaGlnGlnLeuLeu----- 15
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX 1 ATGAGTACAGTCAAGGAAATCAAGCTGCAATAGCTGAGAGGAAAGCTG 60
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX 16 -----GlyCysLeuGlyHis 20
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 AGAGCCCTTGATCCGACAGACTCGAGATACAGAGCTTACAGAGGCTTG 120
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX 21 GlyAlaLeuValLeuGlnLeuSerPheMetLeuAlaGlyLeuValAla 40
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 GGTCCCTGAGTCTGTAACCTCTCTTACGCTCTTGGCTGGC-----CTC 168
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX 41 LeuValGlnValSerLysValProSerSerLeuGlnGlnInsSerGln 60
XX | | | | | | | | | | | | | | | | | | | | | | | |
XX 169 CTTCCTCAACTGCTCAAGGCTGGCAAGCTCAATCAAGTCAAGGAAAGG 228

```

```

QY 61 ILTYGlnAsnLeuThrGlnLeuGlnAlaValAlaGlyLeuLeuSerHis 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 229 ALCTACGAGAACCTGACGACAGCTTAAAGCTGGCATGGCTGACGTCAG 60
QY 81 LeuGlnGlnLeuLeuThrGlnLeuGlnAlaValAlaGlyLeu 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 289 CTGATACAGATCTAAATCAAGTCAAGCTGCAAGCTGCAAGCTGCAAG 60
QY 101 LysSerLysLeuGlnGlnLeuLeuThrGlnLeuGlnAlaValAla 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 349 AATATACAGCTGAGGAAATCTAAAGAGTCAAGCTGCAAGCTGCAAG 60
QY 121 LeuProGlnLeuSerLysLeu----- 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 409 CTTCGAGGAATCTAAGCTGACAGAGATCTACAGAGCTGCAAGCTG 60
QY 127 ----- 40
DB 469 GTGATACAGCTTCAAGAAATCTAAATCAAGTCAAGCTGCAAGCTG 60
QY 128 -----GlnGlnLeuGln 40
DB 529 AAGGCTACAGTCTGAGCTTCTACAGAAATCTAAGTCAAGCTGCAAG 60
QY 135 ThrAlaLeuLysAlaAlaValAlaGlyLeuProGlnLysSerLysLeu 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 589 ACCGCGCTGAAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 60
QY 155 GlnGlnLeuThrAlaLeuLysAlaAlaValAlaGlyLeuProGlnLys 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 649 CAGAGAGCTGAGGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAG 60
QY 175 GlnLeuLeuGlnGlnLeuThrAlaLeuLysAlaAlaValAlaGly 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 709 GAGATCTACAGCTGAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTG 60
QY 195 ProLysAspTrpThrPhePheGlnGlnLysAsnGlySerPheMetSer 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 769 CCGTGGGAAATGCAATCTTCCAGAGGAACTCTTCTTCAATGCTAACT 60
QY 215 TrpHisAspSerValThrAlaCysGlnGlnValAlaAlaGlnLeuVal 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 829 TGGCATACCTCCATCACCGGCGGAGAAAGAAATGGGAGGCTGAGCTG 60
QY 235 AlaGlnGlnGlnAsnPheLeuGlnGlnLeuThrSerArgSerAsnArg 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 889 GCTGATACAGCTGAGCTTCTTCAAGTCAAGCTGCAAGCTGCAAGCT 60
QY 255 GlyLeuSerAspLeuAsnGlnGlnGlnThrTrpAlaGlnValAspLys 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 949 GATCTCTCAAGATCTAAATCAAGGAAATCAAGCTGCAAGCTGCAAG 60
QY 275 ProSerPheGlnAlaGlyTrpAsnSerGlyLysProAsnAsnSerLys 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 1009 CCCAGCTTCAAGCTGAGTATTAATCAAGTCAAGCTGCAAGCTGCAAGCT 60
QY 295 AlaGlnPheSerGlySerGlyTrpAsnAspAsnArgGlyAspAlaLys 40
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 1069 GCGGAAATTTACTGCAATGAGTCAAGGAAATCAAGCTGCAAGCTGCA 60
QY 315 CysLysLysProAlaAla-----GlyHisArgAspGln 325
DB | | | | | | | | | | | | | | | | | | | | | | | |
DB 1129 TGCAAAAAAGGCGGAGGCTTCTGCTCAAGGAGTAA 1164

```

RESULT 3
 AAA88740
 ID AAA88740 standard; cDNA; 1215 bp.
 XX
 AC AAA88740;
 XX
 DT 19-FEB-2001 (first entry)


```

AC AAI65446.1
AD AAI65461 standard; DNA; 1312 bp.
AE
AF
AG AAI65461:
AH
AI 10-DEC-2001 (first entry)
AJ
AK
AL DNA encoding human DC-SIGN, a dendritic cell-specific type I lectin
AM
AN Human receptor, DC-SIGN, dendritic cell; T lymphocyte HIV,
AO gp120, C-type lectin; ICAM3; HIV entry; T cell; macrophage;
AP KW HIV infection; ss
AQ
AR Homo sapiens.
AS
AT Homo sapiens.
AU Location/Qualifiers
AV Key 42..1256
AW CDS /tag=cd
AX /product="DC-SIGN"
AY M0200164752-A2.
AZ
BA
BB 07-SEP-2001.
BC
BD 28-FEB-2001; 2001WD-US063322.
BE
BF 02-MAR-2000; 2000MS-0517605.
BG
BH (UYNV ) UNIV NEW YORK STATE.
BI (UYNV-) UNIV NIMMGEN.
BJ
BK Litzman DR, Kwon D, Van Kooyk Y, Weitenbeck J;
BL
BM WP1; 2001 602565/48
BN P-PSDB; AAC79086.
BO
BP An antibody for the treatment or prevention of HIV-infection comprises
BQ a gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding
BR of DC-SIGN due to concomitant conformational change.
BS
BT
BU Example 1: Page 112-114; 131pp; English.
BV
BW
BX The present sequence encodes a human receptor designated DC-SIGN, that
BY is specifically expressed on dendritic cells and facilitates infection
BZ of T lymphocytes with HIV. DC-SIGN is identical to a HIV-1
CA gp120-binding C-type lectin. DC-SIGN binds ICAM-3 (which is expressed
CB constitutively on T lymphocytes) with high affinity. The specification
CC describes an antibody which is specific for an antigenic fragment of
CD gp120. This antigenic fragment binds to DC-SIGN or is exposed upon
CE gp120 binding of DC-SIGN due to concomitant conformational change.
CF The antibody inhibits the trans enhancement of HIV entry into a T
CG cell or macrophage facilitated by dendritic cells. The antibody is
CH useful to treat or prevent HIV infection.
CI
CJ Sequence 1312 BP; 336 A; 348 G; 372 C; 256 T; 0 other;
CK
CL Alignment Scores:
CM Prod. No.: 1 230-115 Length: 1312
CN Score: 1412.50 Matches: 282
CO Percent Similarity: 77.04% Gaps-valued: 0
CP Best Local Similarity: 71.94% MismatchChes: 19
CQ Query Match: 83.28% Indels: 71
CR DB: Gaps: 4
CS
CT US-09-831-458A-12 (1-325) x AAI65461 (1-1312)
CU
CV 1 MetSerAspSerLysGluIleProArgValGlnGlnIleuGlyLeuLeu ..... 15
CW |||||||
CX |||||||
CY 4.2 ATGGAGTACTACTAAAGGAACAAATCAATCATAGAACATGGGAGAGGAGCAACAAGCTG 101
CZ 16 ..... -GTCGTGCTGCTGTTT 20
DA 102 AGAAGCTTGATTCGCAGACAAAGCTCAAGGATATCAAGAAGCTTAGACAGAGGAT 161

```


[illegible]

OY	249	naahibserftrpmetsglylonsrarspluasoglinglglvthrtprlatrpvalas	269
DB	1964	cggcttccgtcgcattgacattttacagacttaantcagagaacacgcttcgaatgagctgca	2023
OY	2649	pcdysgftrprousrfrbscrftbcstic AcGtyTrrpmsrgly	284
ID	AAH08239		
XX	AAH08239	standard; cDNA; 792 bp.	
AC	AAH08239:		
XX	26-JUN-2001	(first entry)	
DE	Human cDNA clone (5'-primer) SEQ ID NO:5074.		
XX	Human: primer: detection; diagnosis; antisense therapy; gene therapy; ss		
OS	Homo sapiens.		
PN	EPI074617-A2.		
PO	07-FEB-2001.		
PF	29-JUL-2000P 0114126.		
PR	29-JUL-1999J 99JP-0248046.		
PR	27-AUG-1999J 99JP-0300253.		
PR	11-JAN-2000J 2000JP-0118776.		
PR	02-MAY-2000J 2000JP-0184377.		
PR	09-JUN-2000J 2000JP-0241899.		
PA	(HELI-) HELIX RES. INST.		
P1	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;		
P1	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;		
LE	WetL 2001 416733/34.		
TT	Primer sets for synthesizing polynucleotides, particularly the 5602		
PT	full-length cDNAs defined in the specification, and for the detection		
PT	and/or diagnosis of the abnormality of the proteins encoded by the		
PT	full-length cDNAs -		
PS	Claim 1; SEQ ID 5074; 2537bp + CD ROM; English.		
CC	The present invention describes primer sets for synthesising 5602		
CC	full length cDNAs defined in the specification. Where a primer set		
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary		
CC	to the complementary strand of a polynucleotide which comprises one of		
CC	the 5602 nucleotide sequences defined in the specification, where the		
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination		
CC	of an oligonucleotide comprising a sequence complementary to the		
CC	complementary strand of a polynucleotide which comprises a 5' end		
CC	sequence and an oligonucleotide comprising a sequence complementary to a		
CC	polynucleotide which comprises a 3' end sequence, where the		
CC	oligonucleotide comprises at least 15 nucleotides and the combination of		
CC	the 5' end sequence/3' end sequence is selected from those defined in		
CC	the specification. The primer sets can be used in antisense therapy and		
CC	in gene therapy; the primers are useful for synthesising polynucleotides		
CC	particularly full-length cDNAs. The primers are also useful for the		
CC	detection and/or diagnosis of the abnormality of the proteins encoded by		
CC	the full-length cDNAs. The primers allow obtaining of the full-length		
CC	cDNAs easily without any specialised methods. AAH0356 to AAH1528 and		
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to		
CC	AAH15692 represent human amino acid sequences; and AAH15629 to AAH13632		
CC	represent oligonucleotides, all of which are used in the exemplification		
CC	of the present invention.		
XX	Sequence 792 bp; 214 A; 191 C; 234 G; 149 T; 4 other;		

[illegible]

ID AAH9936 standard; cDNA; 433 bp.
 AC AAH99365:
 XX
 DT 16-OCT-2001 (first entry)
 DE Human protein encoding cDNA sequence SEQ ID NO-200.
 KM Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KM anti-inflamatory; antirheumatic; ant arthritic; immunosuppressive;
 KM antibacterial; endocrine; cardiac; central nervous system; vitreous;
 KM anti-HIV; fungicide; antimetabolite; cardiovascular; anti-injury; anaemia;
 KM anti-allergic; haemostatic; pulmonary; antitumor; osteoporosis; eczema;
 KM dermatological; antiallergic; antistimulant; antibacterial; cytostatic;
 KM neuroprotective; antidepressant; motropic; antiparkinsonian; infection;
 KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KM antihypertensive; rheumatoid arthritis; septic shock; pancreatitis;
 KM cardiac dysfunction; neurodegeneration; cardiac anaphylaxis; autoimmune;
 KM genetic disease; haemopoietic disorder; platelet disorder; asthma;
 KM thrombocytopenia; osteoporosis, severe combined immunodeficiency;
 KM allergic rhinitis; diabetes; multiple sclerosis; depression;
 KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KM neurological disorder; ss.
 XX
 XX Homo sapiens.
 OS
 XX
 XX MO200153455-A2.
 PN
 XX
 PD 26-JUL-2001.
 XX
 XX 22-DEC-2000; 2000MO-0535017.
 PP
 XX 23-DEC-1999; 99NS-0471275.
 PR 21-JAN-2000; 2000NS-0488725.
 PR 25-APR-2000; 2000NS-0552317.
 XX
 XX (HSE-) HYSPO INC.
 PA
 XX
 XX Tang YT, Liu C, Drmanac RJ;
 PI
 DR WP1: 2001-457603/49.
 DP P-PSDB; AAM25424.
 XX
 XX
 FT Isolated human polypeptides encoding polypeptides, useful for the
 PT Treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 XX
 XX
 PS Claim 1: Page 387; 1217pp; English.
 XX
 XX *Abstract:* to AAH9936 encode the human proteins given in AAM25425 to
 CC AAM25426, the proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
 CC central nervous system; vitreous; anti-HIV; fungicide; antimetabolite;
 CC cardiovascular; antirheumatic; anti-allergic; haemostatic; pulmonary;
 CC antitumor; osteoporosis; dermatological; antiallergic; antistimulant;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; motropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polypeptides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polypeptides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infectious, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.
 XX
 SO Sequence 433 bp; 110 A; 114 C; 128 G; 81 T; 0 other;
 Alignment Scores:

Prod. No.:	6,060-38	Length:	433
Score:	527.00	Matches:	115
Percent Similarity:	80.95%	Conservative:	4
Best local Similarity:	78.23%	Mismatches:	24
Query Match:	31.07%	Indels:	4
DH:	22	Gaps:	2

US-09-831-458A-12 (1-325) x AAH99365 (1-433)

QY	3	AspSerLysIleuPrArgValGlnGlnLeuGlyLeuLeu-----	15
PD	3	GACTCCAGGAAACCAAGACTGATGATTTGGTCTCTGATGATGAAATGCAAGAG	62
QY	16	-----GlycylSerLeuPheIleuValAla	22
PD	63	CTTCTATTCTGATGACTCTGATGATTAAGATGATTAAGAGTGGTGTATGAGTGGTAAAGTGGG	122
QY	23	IsoValLeuGlnIleuLeuSerPheMetLeuLeuAlaGlyValLeuValAlaIleLeuVal	42
PD	123	CTGGGTGTCACACTCTCTCTCTTCAAGCTCTCTGGCTGGG-----CTCTGTGTC	170
QY	43	GlnValSerLysValProSerSerLeuSerGlnGlnSerGlnAspAlaIleTyr	62
PD	171	CAAGTGTCCAAAGTCTCTCTATCTCAATAAAGTAAAGTAAATCAAGGCAAGACGGCATCTAC	230
QY	63	GlnAsnLeuThrGlnLeuLysAlaAlaValAlaGlyIleuSerGlnLysSerLysLeuGln	82
PD	231	CATACACTTACCCAGCTTAAATATTCAGTGGTGTAGCTTCCAGAGAAATCCCAAGCTTCAG	290
QY	83	GlnIleTyrGlnGlnIleuThrGlnLeuLysAlaAlaValAlaGlnIleuProGlnLysSer	102
PD	291	GAGATCTACGAGAGCTGATCCCAATTTAAGAGCTGATATGGGTGAGCTTCCACAGCAATCT	350
QY	103	LysLeuGlnGlnIleTyrGlnIleuLeuThrArgIleuLysAlaAlaValIleValIlePro	122
PD	351	AAGTTCATAGATATCTAATTAAGATCTAATCTGGTGAAGGTCGACATGGTGAGCTTCCA	410
QY	123	GlnLysSerLysLeuGlnGln	124
PD	411	GAGAAATCTAAGATGAGGAG	431

Search completed: December 7, 2002, 11:26:53
Job time : 296 secs

GenCore version 5.1.4
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW mode)

Run on: December 7, 2002, 00:41:20 ; Search time per sequence
(without alignments)
787.827 Million cell updates/sec

Title: US-09-831-458a-12

Perfect score: 1696

Sequence: 1 MSDSKPRVQCGIIGCTGCH

Scoring table:

Gapop 10 0, Gapext 0 5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

1: SPTRMBL_21: *
2: sp_archaea: *
3: sp_bacteria: *
4: sp_fungi: *
5: sp_human: *
6: sp_invertebrate: *
7: sp_mhc: *
8: sp_organelle: *
9: sp_phage: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_rv1rus: *
16: sp_bacteriaph: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1646.5	97.1	376	4	Q9RXX3
2	1639	96.6	399	4	Q9H2X3
3	1628	96.0	399	4	Q9H2Q9
4	1494.5	88.1	332	4	Q96GP5
5	1412.5	83.3	404	4	Q9RNX6
6	1404.5	82.8	404	6	Q95L98
7	1375	81.1	381	6	Q95L98
8	1375	81.1	381	6	Q95L98
9	1371	80.8	404	4	Q96Q01
10	1369.5	80.7	398	4	Q96Q08
11	1365	79.5	381	6	Q95L98
12	1347.5	77.6	404	6	Q95L98
13	1315.5	77.6	360	4	Q96Q04
14	1314.5	77.8	380	4	Q96Q05
15	1303	75.1	380	4	Q96Q09
16	1274	75.1	380	4	Q96Q00

17	1215.5	71.7	312	4	Q96Q07
18	1157	58.2	296	4	Q9V2A8
19	1148	67.7	263	4	Q96Q04
20	1112.5	65.6	268	4	Q96Q03
21	901	53.1	232	4	Q96Q06
22	874	51.5	297	4	Q96Q08
23	848	50.0	272	4	Q96Q04
24	720	42.5	325	1	Q95L98
25	711.5	42.0	251	4	Q96Q03
26	711.5	42.0	311	1	Q96Q04
27	710	41.9	216	4	Q96Q00
28	631	37.2	295	1	Q95L98
29	612	36.1	208	1	Q912W8
30	588.5	34.7	237	1	Q912W8
31	562	33.1	207	1	Q9V1K4
32	546.5	32.2	238	1	Q912X1
33	538.5	31.8	178	1	Q912W9
34	503.5	29.7	148	4	Q96Q07
35	486.5	28.7	211	1	Q912W5
36	464.5	27.4	168	4	Q96Q06
37	428.5	25.3	274	1	Q96Q09
38	413	24.4	194	1	Q96Q01
39	369	21.8	331	1	Q96Q01
40	351	20.7	309	1	Q96Q07
41	344.5	20.3	309	1	Q96Q03
42	334.5	19.7	309	1	Q96Q02
43	313.5	18.5	293	1	Q96Q08
44	306	18.0	284	1	Q91Y84
45	303.5	17.9	292	4	Q91458

ALIGNMENTS

RESULT 1					
Q9RXX3	PRELIMINARY:	PR1:	376 AA.		
AC	Q9RXX3:				
DT	01-JUN-2001 (TREMBLrel. 17, last sequence update)				
DT	01-JUN-2001 (TREMBLrel. 17, last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 23, last annotation update)				
DE	1-SIGN.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; E				
OC	Mammalia; Eutheria; Primates; Catartida; H				
OX	NCBI_TaxID:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE 21157436, PubMed 1157134;				
RA	Rashirova A A, Gelfandevsk R N, van G				
RA	van Vliet S J, Ellering J B G, Martin M P, Wu L				
RA	Viebig N, Knolle P A, KewalRamani V N, van Kooyk Y,				
RT	*A dendritic cell-specific intercellular adhesion mole				
RT	nonintegrin (DC-SIGN)-related protein is highly expres				
RT	liver sinusoidal endothelial cells and promotes HIV-1				
RT	J. Exp. Med. 193:671-678(2001).				
DR	EMBL: AF290887; AAK20998.1; -				
DR	HSSR: P22897; IE93				
DR	InterPro: IPR002353; Multirep1.				
DR	InterPro: IPR001304; LectinL1.				
DR	Ffam, Pf00055; Lectin_C1.				
DR	PRINTS: PR03356; ANTI-FREDETH1.				
DR	SMART: SM00344; CLEST_1.				
DR	PROSITE: PS00615; CTYPE_LECTIN_1; UNKNOWN_1.				
DR	PROSITE: PS00411; CTYPE_LECTIN_2; -				
SO	SEQUENCE 376 AA; 42724 MW; 36261991RCHE90 CROCA;				
Query Match	97.1%	Score	1646.5	EH 4	141
Best Local Similarity	86.2%	Pred. No.	5.7e-100		
Matches	324	Conservative	1	Mismatches	0
Q9	1 MSDSKPRVQCGIIGCTGCH				

```

Db 1 MSDSKREPVQGLILEDDPTSGIRLPPHDPFOQIHHKSSITGGLGALVLIQILSPML 60
QY 3 LACVVALIIVQSKVPSSLSQFQSDQDAIYQNIYQLKAAGVLSKRSKIQIYQIYQLK 92
Db 61 LACVVALIIVQSKVPSSLSQFQSDQDAIYQNIYQLKAAGVLSKRSKIQIYQIYQLK 120
QY 93 AAVGELPEKSKLOEIYQIYQLKAAGVLSKRSKIQIYQIYQLK----- 136
Db 121 AAVGELPEKSKLOEIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQI 180
QY 137 -----LKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQI 189
Db 181 IYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQI 240
QY 190 LACVVALIIVQSKVPSSLSQFQSDQDAIYQNIYQLKAAGVLSKRSKIQIYQIYQLK 249
Db 241 LACVVALIIVQSKVPSSLSQFQSDQDAIYQNIYQLKAAGVLSKRSKIQIYQIYQLK 300
QY 250 RFSMGLSDINQGTQWQVWVGSPLSPSPQRYWNSQEPNNSQEDCAEFSGSGMNRCQV 309
Db 301 RFSMGLSDINQGTQWQVWVGSPLSPSPQRYWNSQEPNNSQEDCAEFSGSGMNRCQV 360
QY 310 DNYWICKKPAACRDE 325
Db 361 DNYWICKKPAACRDE 376

```

RESULT 2

```

QYH2X3 1D QYH2X3 PRELIMINARY: PRT: 399 AA.
AC QYH2X3
DT 01-MAR-2001 (TREMblrel 16, created)
DT 01-MAR-2001 (TREMblrel 16, last sequence update)
DT 01-JUN-2002 (TREMblrel 21, last annotation update)
DE Probable mannose binding C-type lectin DC-SIGN (MUC-SIGN2 type I
  isoform).
  ABC OR CD209L.
  OS Homo sapiens (Human).
  OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
  OX NCBI_TaxID=9606;
  KN 11
  KE SEQUENCE FROM N.A.
  RX MEDLINE=20472367; PubMed 10975799;
  RA Solileux E.J., Barten R., Trowsdale J.;
  RT "DC-SIGN, a related gene, DC-SIGNR, and CD23 form a cluster on
  19p13."
  RT J. Immunol. 165:2937-2947(2000).
  FI 12
  KN SEQUENCE FROM N.A.
  RP Solileux E.J.;
  RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases
  RN 13
  RX PubMed=1137487;
  KA Mummidi S., Calano G., Lam L., Hoefle A., Telles V., Bequim K.,
  RA Jimenez F., Ahuja S.S., Ahuja S.K.;
  RT "Extensive repertoire of Membrane-bound and soluble Dendritic Cell-
  specific ICM-3-grabbing Nectin-1 (DC-SIGN) and DC-SIGN2
  isoforms: Inter-individual Variation in Expression of DC-SIGN
  transcripts."
  RT J Biol Chem 276:33196-33212(2001)
  DR EMBL: AF209481; AAG1815.2;
  DR EMBL: AF209480; AAG1815.2; JOINED.
  DR EMBL: AF042344; AAK91859.1;
  DR HSSP: P22897; IEQG.
  DR InterPro: IPR002353; AntiFceR2L1.
  DR InterPro: IPR001304; Lectin_C.
  DR Pfam: PF00659; Lectin_C_1
  DR PRINTS: PR00356; ANTI-FCE2EL1.
  DR SMART: SM00034; CLECT_1
  DR SMART: PS00615; C_TYPL1_PFTIN_1; UNKNOWN_1
  DR PROSITE: PS0041; C_TYPL1_PFTIN_2;

```

```

KW Lectin.
SO SEQUENCE 499 AA; 45350 MW; 01A0C9972AFA594 CPG64;

```

```

Query Match 96.6%; Score 1639; DB 4; Length 399;
Post local similarity 81.5%; Pred. No. 1,9e+99;
Matches 325; Conservative 0; Mismatches 0; Indels 74; Gaps 2;

```

```

QY 1 MSDSKREPVQGLILEDDPTSGIRLPPHDPFOQIHHKSSITGGLGALVLIQILSPML 32
Db 1 MSDSKREPVQGLILEDDPTSGIRLPPHDPFOQIHHKSSITGGLGALVLIQILSPML 60
QY 3 LACVVALIIVQSKVPSSLSQFQSDQDAIYQNIYQLKAAGVLSKRSKIQIYQIYQLK 92
Db 61 LACVVALIIVQSKVPSSLSQFQSDQDAIYQNIYQLKAAGVLSKRSKIQIYQIYQLK 120
QY 93 AAVGELPEKSKLOEIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQI 152
Db 121 AAVGELPEKSKLOEIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQI 180
QY 153 IYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQI 166
Db 181 IYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQI 240
QY 167 LPDQSKQQLIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQIYQIYQLK 226
Db 241 LPDQSKQQLIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQIYQIYQLK 300
QY 227 ACIVVLRKIAHGHQIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQIYQIYQLK 286
Db 401 ACIVVLRKIAHGHQIYQIYQLKAAGVLSKRSKIQIYQIYQLKAAGVLSKRSKIQIYQIYQLK 460
QY 287 NSNSINQDAIESSCQNNINQVWQVWVGSPLSPSPQRYWNSQEPNNSQEDCAEFSGSGMNRCQV 325
Db 461 NSNSINQDAIESSCQNNINQVWQVWVGSPLSPSPQRYWNSQEPNNSQEDCAEFSGSGMNRCQV 399

```

RESULT 3

```

QYH209 1D QYH209 PRELIMINARY: PRT: 399 AA.
AC QYH209
DT 01-MAR-2001 (TREMblrel 16, created)
DT 01-MAR-2001 (TREMblrel 16, last sequence update)
DT 01-JUN-2002 (TREMblrel 21, last annotation update)
DE Probable mannose binding C-type lectin DC-SIGNR.
  OS Homo sapiens (Human).
  OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
  OX NCBI_TaxID=9606;
  KN 11
  KE SEQUENCE FROM N.A.
  RX MEDLINE=20472367; PubMed 10975799;
  RA Solileux E.J., Barten R., Trowsdale J.;
  RT "DC-SIGN, a related gene, DC-SIGNR, and CD23 form a cluster on
  19p13."
  RT J. Immunol. 165:2937-2947(2000).
  FI 12
  KN SEQUENCE FROM N.A.
  RP Solileux E.J.;
  RL Submitted (JAN 2001) to the EMBL/GenBank/DBJ databases.
  RN 13
  RX PubMed=1137487;
  KA Mummidi S., Calano G., Lam L., Hoefle A., Telles V., Bequim K.,
  RA Jimenez F., Ahuja S.S., Ahuja S.K.;
  RT "Extensive repertoire of Membrane-bound and soluble Dendritic Cell-
  specific ICM-3-grabbing Nectin-1 (DC-SIGN) and DC-SIGN2
  isoforms: Inter-individual Variation in Expression of DC-SIGN
  transcripts."
  RT J Biol Chem 276:33196-33212(2001)
  DR EMBL: AF209481; AAG1815.2;
  DR EMBL: AF209480; AAG1815.2; JOINED.
  DR EMBL: AF042344; AAK91859.1;
  DR HSSP: P22897; IEQG.
  DR InterPro: IPR002353; AntiFceR2L1.
  DR InterPro: IPR001304; Lectin_C.
  DR Pfam: PF00659; Lectin_C_1
  DR ELMIS: EE0352; ANTI-FCE2EL1.
  DR SMART: SM00034; CLECT_1
  DR SMART: PS00615; C_TYPL1_PFTIN_1;
  DR PROSITE: PS0041; C_TYPL1_PFTIN_2;
  KW Lectin.
  SQ SEQUENCE 499 AA; 45350 MW; FPF0940472AFA594 CPG64;

```

```

Query Match 96.6%; Score 1628; DB 4; Length 399;
Post local similarity 81.2%; Pred. No. 9.8e+99;
Matches 324; Conservative 0; Mismatches 1; Indels 74; Gaps 2;

```

```

UY 1 MSDSKEPPVQOGLT-----GTSHALVAILSPML 32
DB 1 MSDSKEPPVQOGLTLEHPPTTSGIRLPPRDPQODTHCHKSSTTGTGHATVQITSPML 60
UY 33 LAGVAVAILVQVSPSSLSQESQDAIYQNTOLKAAVGELEKSKLOEITQK 92
DB 61 LAGVAVAILVQVSPSSLSQESQDAIYQNTOLKAAVGELEKSKLOEITQK 120
UY 93 AAVGELPEKSKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQK 152
DB 121 AAVGELPEKSKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQK 180
UY 153 IYQELTR-----LKAAGE 166
DB 181 IYQELTRKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQK 240
UY 167 LPDQSKQOQITQVQELTDKTAERKCRKPKDMTFQGNQVFNSSQNMHDSVIAQVEK 226
DB 241 LPDQSKQOQITQVQELTDKTAERKCRKPKDMTFQGNQVFNSSQNMHDSVIAQVEK 300
UY 227 AAVVITKTAEPQNTLQGTSPSNPFSWMTSTLNGPCTWQVWVNSPSPSPQYVWNSGP 286
DB 301 AAVVITKTAEPQNTLQGTSPSNPFSWMTSTLNGPCTWQVWVNSPSPSPQYVWNSGP 360
UY 287 NNSGNEPCAEFSQSGSWNNPCTDVNNTWICKKPAACFDE 325
DB 361 NNSGNEPCAEFSQSGSWNNPCTDVNNTWICKKPAACFDE 399

```

RESULT 4

```

ID 0960P5 PRELIMINARY PRI: 332 AA.
AC 0960P5:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE SDC-SIGN2 type I isoform.
CN CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S., Catalano G., Lam L., Hoeftle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RA *Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms: Inter-Individual Variation in Expression of DC-SIGN
RT Transcripts.*;
RL J. Biol. Chem. 276:33196-33212(2001).
DB EMBL: AY042238; AAK91863.1;
DB InterPro: IPR001304; Lectin_C1;
DB Pfam: PF00059; Lectin_C1;
DB PROSITE: PS00615; G-TYPE_LLECTIN_1; UNKNOWN_1;
DB PROSITE: PS00615; G-TYPE_LLECTIN_2; 1;
DE PROSITE: PS00615; G-TYPE_LLECTIN_2; 1;
SEQUENCE 332 AA: 37953 MW: 40769.47 Da; pI: 4.74;

```

```

Query Match 88.1% Score 1494.5 PR 4: Length 332;
Best Local Similarity 85.2% Pred. No. 41e-90;
Matches 293; Conservative 5; Mismatches 15; Indels 31; Gaps 2.
UY 1 MSDSKEPPVQOGLT-----GCTGHGAVIQLISPMILAVAIL 41
DB 1 MSDSKEPPVQOGLTLEHPPTTSGIRLPPRDPQODTHCHKSSTTGTGHATVQITSPML 48
UY 42 VQVSPSSLSQESQDAIYQNTOLKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQK 101
DB 49 VQVSPSSLSQESQDAIYQNTOLKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQK 108
UY 102 SKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQK 161

```

```

DB 109 SKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQK 166
UY 162 AAVGELPEKSKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQK 226
DB 169 AAVGELPEKSKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQKTPKAAVGELEKSKLOEITQK 226
UY 222 GQVFAVAVITKTAEPQNTLQGTSPSNPFSWMTSTLNGPCTWQVWVNSPSPSPQYVWNSGP 286
DB 229 GQVFAVAVITKTAEPQNTLQGTSPSNPFSWMTSTLNGPCTWQVWVNSPSPSPQYVWNSGP 300
UY 287 NNSGNEPCAEFSQSGSWNNPCTDVNNTWICKKPAACFDE 325
DB 289 NNSGNEPCAEFSQSGSWNNPCTDVNNTWICKKPAACFDE 392

```

RESULT 5

```

ID 09NNX6 PRELIMINARY PRI: 404 AA.
AC 09NNX6:
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Membrane-associated lectin type-C (probably mouse hind
DE lectin DC-SIGN) (MDC-SIGNIA type I isoform).
CN CD209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=PLACENTA;
RX MEDLINE=92390446; PubMed=1518859;
RA Curtis B.M., Scharnowski S., Watson A.J.;
RA *Sequence and expression of a membrane-associated C-type
RT exhibits CD4-independent binding of human immunodeficient
RT envelope glycoprotein gp 120.*;
RL Proc Natl Acad Sci U S A 89:4456-4460(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20432467; PubMed=10975799;
RA Solleux E.J., Barten R., Trowsdale J.;
RA *DC-SIGN, a related gene, DC SIGNR, and CD24 form a cluster
RT 19p13.*;
RL J. Immunol. 165:2937-2942(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21157496; PubMed=11257144;
RA Bashirova A.A., Geiltenbeck T.K.H., van Dalenhoven G.C.J.,
RA van Vliet S.T., Flierberg R.G., Martin M.P., Wu L., Mar
RA Vlieghe N., Koelle P.A., Kwak-Kwami V.N., van Kooyk Y.,
RA *A dendritic cell specific intercellular adhesion molecule
RT nonintegrin (DC-SIGN)-related protein is highly expressed
RT in liver sinusoidal endothelial cells and promotes HIV-1 in
RL J. Exp. Med. 193:671-678(2001).
RN [4]
RP SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S., Catalano G., Lam L., Hoeftle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RA *Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms: Inter-Individual Variation in Expression of DC-SIGN
RT Transcripts.*;
RL J. Biol. Chem. 276:33196-33212(2001).
DB EMBL: M98457; AF77072.1;
DB EMBL: AF204479; AAG13814.1;
DB EMBL: AF290886; AAK20997.1;
DB EMBL: AY042221; AAK91846.1;
DB HSSP: P32897; IECG;
DB InterPro: IPR001304; Lectin_C1;
DB Pfam: PF00059; Lectin_C1;

```

SMART: SM00034; CLPCT: 1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00641; C_TYPE_LECTIN_2; 1.
 KW Lectin.
 SO SEQUENCE 404 AA; 45775 MW; A23FA260145340V CIRC64;

Query Match 83.38; Score 1412.5; DB 4; Length 404;
 Best Local Similarity 71.98; Prod. No. 120-84;
 Matches 282; Conservative 20; Mismatches 19; Indels 71; Gaps 4;

QY 1 MSDSKEPRVQGLL-----GCLGHALVQLLSFMLACVLVAI 40
 DB 1 MSDSKEPRVQGLL-----GCLGHALVQLLSFMLACVLVAI 40
 QY 41 LVVSKVPSSTLSQPSQSDAIYQNLTLQKAAGELSEKSKLOEYQELTLQKAAGELPE 100
 DB 57 LVVSKVPSSTLSQPSQSDAIYQNLTLQKAAGELSEKSKLOEYQELTLQKAAGELPE 116
 QY 101 KSKLOEYQELTLQKAAGELPEKSKL-----127
 DB 117 KSKLOEYQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPEKSKLOEYQELTL 176
 QY 128 -----OEYQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPUSKQ 174
 DB 177 KAAGELPEKSKLOEYQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPUSKQ 236
 QY 175 QYVQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPUSKQ 234
 DB 237 EYVQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPUSKQ 296
 QY 235 AEQNFLOQTSNSNPTWNGI-SH-NQSTWLVWVDSPLSPSPQYWNNGPNNNGNEK 294
 DB 297 AEQNFLOQTSNSNPTWNGI-SH-NQSTWLVWVDSPLSPSPQYWNNGPNNNGNEK 356
 QY 295 AEQNSGMDNPGVQVNWYCKKPAK-CPEPE 325
 DB 357 AEQNSGMDNPGVQVNWYCKKPAK-CPEPE 388

RESULT 6

Q95L98 PRELIMINARY; PRI: 404 AA.
 AC Q95L98;
 DI 01-DEC-2001 (TEMBUREL 19, Created)
 DI 01-DEC-2001 (TEMBUREL 19, Last sequence update)
 DI 01-JUN-2002 (TEMBUREL 21, Last annotation update)
 DE Dendritic cell-specific ICAM-3 grabbing non-integrin (Fragment).
 OS DC-SIGN.
 OC Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan
 ON NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Geil, J. H., Koyama, G., van Halbeek, H., van Vliet, A. J.,
 van Schijndel, A. C., Engering, A., Heeney, J. L., van Kooyk, Y.,
 "Identification of Rhesus and chimpanzee DC-SIGN, both act as HIV/SIV
 gp120 trans-receptors similar as human DC-SIGN."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF191085; AK97458.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C_1
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00641; C_TYPE_LECTIN_2; 1.
 KW Integrin.
 FT NON_FER 404
 SO SEQUENCE 404 AA; 45885 MW; HFD872D38B48CFE CIRC64;

Query Match 82.88; Score 1404.5; DB 6; Length 404;
 Best Local Similarity 70.98; Prod. No. 38e-84;
 Matches 279; Conservative 23; Mismatches 20; Indels 71; Gaps 4;
 QY 1 MSDSKEPRVQGLL-----GCLGHALVQLLSFMLACVLVAI 40

DB 1 MSDSKEPRVQGLL-----GCLGHALVQLLSFMLACVLVAI 40
 QY 41 LVVSKVPSSTLSQPSQSDAIYQNLTLQKAAGELSEKSKLOEYQELTLQKAAGELPE 100
 DB 57 LVVSKVPSSTLSQPSQSDAIYQNLTLQKAAGELSEKSKLOEYQELTLQKAAGELPE 116
 QY 101 KSKLOEYQELTLQKAAGELPEKSKL-----150
 DB 117 KSKLOEYQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPEKSKLOEYQELTL 176
 QY 151 -----OEYQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPUSKQ 174
 DB 177 KAAGELPEKSKLOEYQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPUSKQ 236
 QY 175 QYVQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPUSKQ 234
 DB 237 EYVQELTLQKAAGELPEKSKLOEYQELTLQKAAGELPUSKQ 296
 QY 235 AEQNFLOQTSNSNPTWNGI-SH-NQSTWLVWVDSPLSPSPQYWNNGPNNNGNEK 294
 DB 297 AEQNFLOQTSNSNPTWNGI-SH-NQSTWLVWVDSPLSPSPQYWNNGPNNNGNEK 356
 QY 295 AEQNSGMDNPGVQVNWYCKKPAK-CPEPE 325
 DB 357 AEQNSGMDNPGVQVNWYCKKPAK-CPEPE 388

RESULT 7

Q95L98 PRELIMINARY; PRI: 381 AA.
 AC Q95L98;
 DI 01-DEC-2001 (TEMBUREL 19, Created)
 DI 01-DEC-2001 (TEMBUREL 19, Last sequence update)
 DI 01-MAR-2002 (TEMBUREL 26, Last annotation update)
 DE Dendritic cell-specific ICAM-3 grabbing non-integrin.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 ON NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE 2145051, PubMed-11581396;
 RA Barilaud F., Pohlmann S., Sharwasser T., Kimata M.T., Choi Y.K.,
 Haegarty B.S., Ahmad N., Macfarlan T., Edwards T.G., Leslie G.J.,
 Arnason J., Reinhardt T.A., Kimata J.T., Littman D.R., Hoxie J.A.,
 Doms R.W.,
 "Functional and antigenic characterization of human, Rhesus macaque,
 RT p140 and murine dc-sim."
 RL J Virol 76:10241-10249(2002).
 DR EMBL: AF169755; AML14438.1;
 DR InterPro: IPR001404; Lectin_C.
 DR Pfam: PF00059; Lectin_C_1
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00641; C_TYPE_LECTIN_2; 1.
 KW Integrin.
 SO SEQUENCE 481 AA; 42897 MW; Q1FE7B3B42C91D49 CIRC64;

Query Match 81.18; Score 1375; DB 6; Length 381;
 Best Local Similarity 74.08; Prod. No. 3e-82;
 Matches 275; Conservative 25; Mismatches 25; Indels 48; Gaps 4;
 QY 1 MSDSKEPRVQGLL-----GCLGHALVQLLSFMLACVLVAI 40
 DB 1 MSDSKEPRVQGLL-----GCLGHALVQLLSFMLACVLVAI 56
 QY 41 LVVSKVPSSTLSQPSQSDAIYQNLTLQKAAGELSEKSKLOEYQELTLQKAAGELPE 100
 DB 57 LVVSKVPSSTLSQPSQSDAIYQNLTLQKAAGELSEKSKLOEYQELTLQKAAGELPE 116
 QY 101 KSKLOEYQELTLQKAAGELPEKSKL-----160


```

DB 117 KSKQGFIVPEITLTKAAVGLPEKSKIOEIVQELTELKAAVGLPEKSKIOEIVQDLSPL 176
UY 161 KAAGE-----LPDQSKQOUIYQELFDEKTAPELDRKCPMD 197
DB 177 KAAGVDLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQDLSPL 256
UY 198 WTEFGNCFYFMSNSORNNHDSVTAQCEVKAJLVV:KIAFAUNLVLQV:SNKNT:SNKGLS 257
DB 237 WTEFGNCFYFMSNSORNNHDSVTAQCEVKAJLVV:KIAFAUNLVLQV:SNKNT:SNKGLS 256
UY 258 DLNCGTQWQVDSPLSPFORYNNSGPNNSGNDCAEFGSGGWNDRCDVQNTCKK 317
DB 297 DLNCGTQWQVDSPLSPFORYNNSGPNNSGNDCAEFGSGGWNDRCDVQNTCKK 356
UY 418 FVA-CFPRD 325
DB 457 SAASCODE 365

RESULT 8
DB 096008 PRELIMINARY: PRT: 381 AA.
AC 096008:
DT 01-JUN-2002 (TREMblrel, 21, Created)
DT 01-JUN-2002 (TREMblrel, 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DB Type 11 membrane protein CD209.
DB Macaca mulatta (Rhesus macaque).
DB Eukaryota; Metazoa; Chordata; Clariata; Vertebrata; Euteleostomi;
DB Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
DB Cercopithecoidea; Macaca.
DB NCBI_TaxID=9544;
DB RN [1]
DB RP SEQUENCE FROM N.A.
DB RX MEDLINE=21819420; PubMed=11818574,
DB Wu L., Bashitova A.A., Martin T.D., Villamide J., Mehjoub F.,
DB Chertov A.O., Unutma D., Pope M., Carrington M., Kewakhamri V.N.;
DB "Rhesus macaque dendritic cells efficiently transmit primate
DB RT lentiviruses independently of DC-SIGN";
DB RT Proc. Natl. Acad. Sci. U.S.A. 99:1568-1573(2002).
DB RL EMBL: AY040319; AAK74185.1;
DB DR PROSITE: 481 AAF: 42955 MW: 073278042071043 GRC64,
DB SEQUENCE: 481 AAF: 42955 MW: 073278042071043 GRC64,
DB Query Match 81.1%; Score 1375; DB 6; Length 381,
DB Best Local Similarity 74.0%; Pred. No. 38-82;
DB Matches 273; Conservative 23; Mismatches 25; Indels 48; Gaps 4;

UY 1 MSDSKPRVVOQLGL-----GCLGICGALVQLLSFMLAGVLVAI 40
DB 1 MSDSKPRVVOQLGL-----GCLGICGALVQLLSFMLAGVLVAI 40
UY 41 LVQVSKVSSLSQEGSEDDATAYQNTQLKAAVGLSEKSKIOEIVQELTKAAVGLPE 100
DB 57 LVQVSKVSSLSQEGSEDDATAYQNTQLKAAVGLSEKSKIOEIVQELTKAAVGLPE 116
UY 101 KSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKL 160
DB 117 KSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKL 176
UY 161 KAAGE-----LPDQSKQOUIYQELFDEKTAPELDRKCPMD 197
DB 177 KAAGVDLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQDLSPL 256
UY 198 WTEFGNCFYFMSNSORNNHDSVTAQCEVKAJLVV:KIAFAUNLVLQV:SNKNT:SNKGLS 257
DB 237 WTEFGNCFYFMSNSORNNHDSVTAQCEVKAJLVV:KIAFAUNLVLQV:SNKNT:SNKGLS 256
UY 258 DLNCGTQWQVDSPLSPFORYNNSGPNNSGNDCAEFGSGGWNDRCDVQNTCKK 317
DB 297 DLNCGTQWQVDSPLSPFORYNNSGPNNSGNDCAEFGSGGWNDRCDVQNTCKK 356
UY 418 FVA-CFPRD 325
DB 457 SAASCODE 365

```

```

DB 357 SAASCODE 365

RESULT 9
DB 096001 PRELIMINARY: PRT: 404 AA.
AC 096001:
DT 01-DEC-2001 (TREMblrel, 19, Created)
DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel, 20, Last annotation update)
DB MDC-SIGNIB type I isoform.
DB CD209.
DB Homo sapiens (Human).
DB Eukaryota; Metazoa; Chordata; Clariata; Vertebrata; Eu-
DB Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Ho-
DB NCBI_TaxID=9606;
DB RN [1]
DB RP SEQUENCE FROM N.A.
DB RX PubMed=1137487;
DB FA Mummidi S., Catano G., Lam L., Hoelle A., Jelles V., Le-
DB Jimenez F., Ahuja S.K.;
DB "Extensive Repertoire of Membrane-bound and Soluble Ig-
DB RT specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and
DB RT Isoforms: Inter-Individual Variation in Expression of
DB RT Transcripts";
DB RL J. Biol. Chem. 276:33196-33212(2001).
DB DR EMBL: AY042229; AAK91854.1;
DB DR InterPro: IPR001304; LeclIn-C;
DB DR Pfam: PF00456; IprIn-C; 1;
DB DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DB DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DB SEQUENCE: 404 AAF: 45570 MW: 449570727090030 GRC64,
DB Query Match 80.8%; Score 1371; DB 4; Length
DB Best Local Similarity 72.2%; Pred. No. 5,96 82;
DB Matches 273; Conservative 23; Mismatches 23; Indels
DB 4 SKEPVQVQIG-----LLOGLGICGALVQLLSFMLAGVLVAI 44
DB 15 SEEBQIRGIGFPGTQGYKSLAGTDLHGLVQLLSFLLAG---LTVY 26
UY 55 USKQALVQNTQLKAAVGLSEKSKIOEIVQELTKAAVGLPEKSKIOEIVQDLSPL 114
DB 71 USKQALVQNTQLKAAVGLSEKSKIOEIVQELTKAAVGLPEKSKIOEIVQDLSPL 130
UY 115 KAAGVELPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQDLSPL 176
DB 131 KAAGVELPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQDLSPL 256
UY 129 EIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQDLSPL 317
DB 191 EIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQELTKAAVGLPEKSKIOEIVQDLSPL 356
UY 189 RLCRHCPKMDWTFPGNCFYFMSNSORNNHDSVTAQCEVKAJLVV:KIAFAUNLVLQV:SNKNT:SNKGLS 438
DB 251 RLCRHCPKMDWTFPGNCFYFMSNSORNNHDSVTAQCEVKAJLVV:KIAFAUNLVLQV:SNKNT:SNKGLS 476
UY 249 NPEFMMGLSTNDRPTQWQVDSPLSPFORYNNSGPNNSGNDCAEFGSGGWNDRCDVQNTCKK 516
DB 311 NPEFMMGLSTNDRPTQWQVDSPLSPFORYNNSGPNNSGNDCAEFGSGGWNDRCDVQNTCKK 554
UY 309 VQNYWICKKFAA-CFPRD 325
DB 371 LAKFWICKKSAASCODE 365

```

```

RESULT 10
DB 096008 PRELIMINARY: PRT: 398 AA.
AC 096008:
DT 01-DEC-2001 (TREMblrel, 19, Created)
DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel, 20, Last annotation update)

```

MD-STGNIA type II isoform.
 G6209.
 OS Homo sapiens (Human).
 EC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 M1 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NC NCB1:taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed:11337487;
 KA Minnich S., Catano G., Lam L., Hoefle A., Telles V., Hoegum K.,
 RJ Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-STG1) and DC-STG2
 RT Isoforms: Inter-individual Variation in Expression of DC-STG1
 RT Transcripts";
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042222; AAK91847.1; C.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2.1.
 SO SEQUENCE 398 AA; 45031 MW; D92576486701D2E CRC64;

Query Match 80.78; Score 1369.5; DB 4; Length 398;
 Best Local Similarity 70.4%; Pval No. 7.3e-82;
 Matches 276; Conservative 20; Mismatches 19; Indels 77; Gaps 5;

QY 1 MSDSKEPVOQIGL-----GCLGHGALVLDLSPMLAGVIAI 40
 Db 1 MSDSKEPVOQIGLLEEEGLGLGFRGTGKSLAGLGHGALVLDLSTLLAG-----L 56
 QY 41 LVQVSKVPSSI.SQPSQSDAIVNTQLKAAGVLSKSLQEIYQELTOLKAAGVLE 100
 Db 57 LVQVSKVPSSI.SQPSQSDAIVNTQLKAAGVLSKSLQEIYQELTOLKAAGVLE 116
 QY 101 KSKLOEIYQELTRKAAGVLEPKSKL-----127
 Db 117 KSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTOLKAAGVLEPKSKLOEIYQELTR 176
 QY 128 -----OEIYQELTRKAAGVLEPKSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTR 174
 Db 177 KAAGVLEPKSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTR 246
 QY 175 QIYQELTOLKTAFFRLCRHCPKMTFFQGNCTYMSNSQPMHESVTAQCEVAGLVIAI 244
 Db 237 ELYQELTOLKKAAGVLEPKSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTR 296
 QY 245 AEEVNFLOLQTSRNSPFWMMISPI.NQEGTQWQVWVSPLSFSPQVWNSQEPNNNGEIP 294
 Db 297 AEE-----QSSPSNPFVMMISPI.NQPTWQVWVSPILPSPKQVWNSQEPNNNGEIP 350
 QY 295 AEFSGSNWNNPCTVENVYICKPIAA-CFIDE 325
 Db 351 AEFSGSNWNNPCTVENVYICKPIAA-CFIDE 382

RESULTS 11
 Q95J96 PRELIMINARY; PRT; 381 AA.
 AC Q95J96;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DT Dendritic cell-specific ICAM-3 grabbing nonintegrin.
 OS Macaca mulatta (Pig tailed macaque)
 EC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 M1 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 NC NCB1:taxid=9545;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE:21465051; PubMed:11581196;
 RA Haribaud F., Palmann S., Spatwasser T., Kimata M.T., Choi Y.K.,

Haghighi R.S., Ahmad N., Macfarlan T., Edwards T.G., Leslie G.J.,
 RA Atkinson J., Reinhardt T.A., Kimata J.T., Littman D.R., Hoxie J.A.,
 RA Davis R.W.;
 RT "Functional and anti-pathogen characterization of human, rhesus macaque,
 RT pig-tailed macaque, and murine dc-sign";
 RL J. Virol. 75:10281-10289(2001).
 DR EMBL: AF343727; AAL14428.1; C.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2.1.
 SO SEQUENCE 491 AA; 42951 MW; CAF6E2E3D454B74A CRC64;

Query Match 80.58; Score 1365; DB 6; Length 381;
 Best Local Similarity 71.4%; Pval No. 1.4e-81;
 Matches 271; Conservative 23; Mismatches 27; Indels 48; Gaps 4;

QY 1 MSDSKEPVOQIGL-----GCLGHGALVLDLSPMLAGVIAI 40
 Db 1 MSDSKEPVOQIGLLEEEGLGLGFRGTGKSLAGLGHGALVLDLSTLLAG-----L 56
 QY 41 LVQVSKVPSSI.SQPSQSDAIVNTQLKAAGVLSKSLQEIYQELTOLKAAGVLE 100
 Db 57 LVQVSKVPSSI.SQPSQSDAIVNTQLKAAGVLSKSLQEIYQELTOLKAAGVLE 116
 QY 101 KSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTR 160
 Db 117 KSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTOLKAAGVLEPKSKLOEIYQELTR 176
 QY 161 KAAGV-----LPQSKQOQIYQELTOLKTAFFRLCRHCPK 197
 Db 177 KAAGVLEPKSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTRKAAGVLEPKSKLOEIYQELTR 236
 QY 198 WDFQGNCTYMSNSQPMHESVTAQCEVAGLVIAI.TAPQNTQI.QSSPSNPFVMMIS 257
 Db 237 WDFQGNCTYMSNSQPMHESVTAQCEVAGLVIAI.SAEQNTQI.QSSPSNPFVMMIS 296
 QY 258 DFNQGTQWQVWVSPLSFSPQVWNSQEPNNNGEIPNNQDCAEFSQSNWNNPCTVENVYICK 317
 Db 297 DFNQGTQWQVWVSPLSFSPQVWNSQEPNNNGEIPNNQDCAEFSQSNWNNPCTVENVYICK 356
 QY 318 PAA-CFIDE 325
 Db 357 SAASCSIDE 365

RESULTS 12
 Q95J96 PRELIMINARY; PRT; 404 AA.
 AC Q95J96;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DT Dendritic cell-specific ICAM-3 grabbing non integrin.
 OS Macaca mulatta (Rhesus macaque).
 EC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 M1 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 NC NCB1:taxid=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE: LYMPH NODE;
 AC Gelfand-Jones R.H., Koopman A., van den Hulstwen G.C.F., van Vliet S.J.,
 RA van Schijndel A.C., Luchters A., van der Vliet A., van Kesteren Y.,
 RT "Identification of human and chimpanzee DC-STG1, both act as HIV/SIV
 RT gp120 trans-receptors similar as human DC-STG1";
 RL Submitted (Jun-2001) to the EMBL/GenBank/DDO databases.
 DR EMBL: AF491086; AAK97459.1; C.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.

```

DB PROSITE: PS50041: C_TYPE_LLECTIN_2; 1.
KM Integrin.
SO SEQUENCE 404 AA: 45655 MW: F77953413430CC17 CRG64;

Query Match 79.5%; Score 1347.5; DB 6; Length 404;
Best Local Similarity 68.9%; Pred. No. 20-80;
Matches 270; Conservative 24; Mismatches 27; Indels 71; Gaps 4.

QY 1 MSDSKPRVQDLGLL-----GCLGHALVLDLSTFMLAGVLA1 40
DB 1 MSDSKPRVQDLGLL-----GCLGHALVLDLSTFMLAGVLA1 40
QY 41 LVOYSKVPSSLSQSFQDAIYQNTOLKAAVGLSEKSKLOEYQELTOLKAAVELPE 100
DB 57 LVOYSKVPSSLSQSFQDAIYQNTOLKAAVGLSEKSKLOEYQELTOLKAAVELPE 116
QY 101 KSKLOEYQELTOLKAAVGLSEKSKLOEYQELTOLKAAVGLSEKSKLOEYQELTOL 160
DB 117 KSKLOEYQELTOLKAAVGLSEKSKLOEYQELTOLKAAVGLSEKSKLOEYQELTOL 176
QY 161 KAANGELPDGSKQOQIYQELT-----181
DB 177 KAANGELPDGSKQOQIYQELTOLKAAVGLSEKSKLOEYQELTOLKAAVGLPEPKQ 236
QY 182 -----DLKTAPEELCHCHKDMTFPGNGYFMSNSQRMHDSVTACQEVPAQLVIT 234
DB 237 ELYQELLQTLKAAVGLCHPCWMENTFFQNGYFMSNSQRMHDSVTACQEVPAQLVIT 296
QY 235 AEEQNFLOLQTSRSPNSPMWNTSLNDEFTWQWNTSPISPFQYVNSGEPNNSNEDC 294
DB 297 AEEQNFLOLQTSRSPNSPMWNTSLNDEFTWQWNTSPISPFQYVNSGEPNNSNEDC 356
QY 295 AEEQNFLOLQTSRSPNSPMWNTSLNDEFTWQWNTSPISPFQYVNSGEPNNSNEDC 325
DB 357 AEEQNFLOLQTSRSPNSPMWNTSLNDEFTWQWNTSPISPFQYVNSGEPNNSNEDC 388

RESULT 13
Q96004 PRELIMINARY: PRT: 360 AA.
AC Q96004: 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE SDC-SIGNIA type II isoform.
GN CD209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX PubMed=1137487;
KA Mummidi S., Catano G., Lam L., Hoeft A., Telles V., Begum K.,
KA Jimenez F., Ahuja S.S., Ahuja S.K.;
KA "Extensive repertoire of Membrane-bound and Soluble dendritic Cell-
RT specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
RT isoforms. Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
RL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042226; AAK91851.1;
DR InterPro: IPR001304; Lectin_C;
DR Pfam: PF00059; Lectin_C;
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
SO SEQUENCE 360 AA: 41009 MW: 6ABE299AEDFAA8 CRG64;

Query Match 77.6%; Score 1315.5; DB 4; Length 360;
Best Local Similarity 70.2%; Pred. No. 2-78;
Matches 261; Conservative 19; Mismatches 17; Indels 75; Gaps 3;

QY 1 MSDSKPRVQDLGLLGLGHALVLDLSTFMLAGVLAIVVSVSSLSQSFQDAI 60
DB 1 MSDSKPRVQDLGLLGLGHALVLDLSTFMLAGVLAIVVSVSSLSQSFQDAI 60

```

```

DB 1 MSDSKPRVQDLGLL-----VSKVSS
QY 61 IYQNTOLKAAVGLSEKSKLOEYQELTOLKAAVGLSEKSKLOEYQELTOLKAAVGL 120
DB 33 IYQNTOLKAAVGLSEKSKLOEYQELTOLKAAVGLSEKSKLOEYQELTOLKAAVGL 120
QY 121 LPEKSKL-----
DB 93 LPEKSKLQFIYQELTOLKAAVGLPEKSKLOEYQELTOLKAAVGLPEKSKL 120
QY 135 TRKAAVGLPEKSKLOEYQELTOLKAAVGLPEKSKLOEYQELTOLKAAVGLPEKSK 120
DB 153 TRKAAVGLPEKSKLOEYQELTOLKAAVGLPEKSKLOEYQELTOLKAAVGLPEKSK 120
QY 195 PKDWTFFQNGYFMSNSQRMHDSVTACQEVPAQLVITKAEQNFLOLQTSRSPNS 294
DB 213 PKDWTFFQNGYFMSNSQRMHDSVTACQEVPAQLVITKAEQNFLOLQTSRSPNS 294
QY 255 GLSIQNLPGTWQWNTSPISPFQYVNSGEPNNSNEDC 325
DB 273 GLSIQNLPGTWQWNTSPISPFQYVNSGEPNNSNEDC 325
QY 315 CKKPAV-CFRDE 325
DB 333 CKKSAASGRDE 344

RESULT 14
Q96005 PRELIMINARY: PRT: 480 AA.
AC Q96005: 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE SDC-SIGNIA type I isoform.
GN CD209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX PubMed=1137487;
KA Mummidi S., Catano G., Lam L., Hoeft A., Telles V., Begum K.,
KA Jimenez F., Ahuja S.S., Ahuja S.K.;
KA "Extensive repertoire of Membrane-bound and Soluble dendritic Cell-
RT specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
RT isoforms. Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
RL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042225; AAK91850.1;
DR InterPro: IPR001304; Lectin_C;
DR Pfam: PF00059; Lectin_C;
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
SO SEQUENCE 380 AA: 43330 MW: 6CFWV45ABBEF02 CRG64;

Query Match 77.5%; Score 1314.5; DB 4; Length 380;
Best Local Similarity 71.2%; Pred. No. 2-78;
Matches 263; Conservative 21; Mismatches 31; Indels 74;

QY 1 MSDSKPRVQDLGLLGLGHALVLDLSTFMLAGVLAIVVSVSSLSQSFQDAI 60
DB 1 MSDSKPRVQDLGLL-----EEELRGIGLPGQIRGKRV-----SLAVSKVPSSLS 60
QY 61 IYQNTOLKAAVGLSEKSKLOEYQELTOLKAAVGLPEKSKLOEYQELTOLKAAVGLPEKSK 120
DB 53 IYQNTOLKAAVGLSEKSKLOEYQELTOLKAAVGLPEKSKLOEYQELTOLKAAVGLPEKSK 120
QY 121 LPEKSKL-----
DB 113 LPEKSKLQFIYQELTOLKAAVGLPEKSKLOEYQELTOLKAAVGLPEKSKL 120

```

```

QY 135 TPLKAAGVLPKSKTQETIYQELTPLKAAGVLPPOSKQQTQYQELTTPKTAFFPLCPHC 194
DB 173 TLKKAAGVLPKSKQETIYQELTLKKAAGVLPKSKQETIYQELTLKKAAGVLPCLCPHC 232
QY 195 PKTWTFQGNWYMSNSQPMWHSVLAQGEVPAQLVLIKTAEPNPLQLOTSPNSPSSWM 254
DB 233 PWMTTFQGNWYMSNSQPMWHSITACEKGAQGLVLIKSAEONPLQLOSSRSNRFTWM 292
QY 255 GLSDLNQEGTQWQVQSGPISPSFORWNSGFINNSGNEDECAFSGSGWNNHCDVDWTM 314
DB 293 GLSDLNQEGTQWQVQSGPISPSFORWNSGFINNSGNEDECAFSGSGWNNHCDVDWTM 352
QY 415 CKKPAA-CFRDE 325
DB 453 CKKSAASGRDE 364

```

```

RESULT 15
Q960P9 PRELIMINARY: PRT: 334 AA.
AC Q960P9;
DI 01-DEC-2001 (TREMblrel: 19, Created)
DI 01-DEC-2001 (TREMblrel: 19, Last sequence update)
DI 01-MAR-2002 (TREMblrel: 20, Last annotation update)
DE SDC-SIGNB type II isoform.
GN 30209.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 111
RP SEQUENCE FROM N.A.
RX PubMed-11337487;
KA Mumental S., Catalan G., Lam I., Hoofnagle A., Telles V., Bertram K.,
KA Jimenez F., Abuja S., Abuja S.K.;
KI "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
KI Specific ICAM-3-grabbing NonIntegrin 1 (DC-SIGN) and DC-SIGN2
KI Isoforms: Inter-individual Variation in Expression of DC-SIGN
KI Transcripts.";
PL J. Biol. Chem. 276:33196-33212(2001).
DR EMBL: AY042231; AAK91856.1; -.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF000059; Lectin_C_1.
DR PROSITE: PS00615; G_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; G_TYPE_LLECTIN_2; 1.
SV SEQUENCE 334 AA: 37843 MW: 86796RC4111C86BP CMC64;

```

Query Match 76.8%; Score 1003; DB 4; Length 334;
 Rest Local Similarity 78.0%; Pred. No. 1,3e-77;
 Matches 252; Conservative 25; Mismatches 26; Indels 20; Gaps 2;

```

QY 4 SKKFWQDLGLDGLGHGALVLDLSPMLAGVLAILEVQSKVPSLSQEHSPQDATTQ 63
DB 15 SFEQRLGLGFGHGTGRYKSLA-----VSKVPSISQEHSPQDATTQ 55
QY 64 NLTQLKAAGVLPKSKLQETIYQELTQLKAAGVLPKSKLQETIYQELTLKKAAGVLP 123
DB 56 NLTQLKAAGVLPKSKLQETIYQELTQLKAAGVLPKSKLQETIYQELTLKKAAGVLP 115
QY 124 KSKLQETIYQELTLKKAAGVLPKSKLQETIYQELTLKKAAGVLPPOSKQQTQYQELTDL 183
DB 116 KSKLQETIYQELTLKKAAGVLPKSKLQETIYQELTLKKAAGVLPKSKQETIYQELTOL 175
QY 184 KLAEPICPHCPKDWTFPGQNCYEMKSNQNMHDSVTACQEVPAQLVLIKTAEPNPLQ 243
DB 176 KAAVEPLCHCPKDWTFPGQNCYEMKSNQNMHDSITACEKGAQGLVLIKSAEONPLQ 235
QY 244 QTSRSNRFSSMGLSDLNQEGTQWQVQSGPISPSFORWNSGFINNSGNEDECAFSGSGW 303
DB 236 QSSRSNRFSSMGLSDLNQEGTQWQVQSGPISPSFORWNSGFINNSGNEDECAFSGSGW 295
QY 404 INKQVDNWNWICKKPAA-CFRDE 325
DB 453 CKKSAASGRDE 364

```

DB 296 DKCYNLAKTWCKKSAASGRDE 318

Search completed: December 7, 2002, 11:05:43
 Job Time: 91 secs


```

1 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
2 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
3 US-09-864-761-7350
4
5 Query Match          9.78; Score 153.2; DB 10; Length 592;
6 Best local similarity 98.1%; Prod No 6,9e+6;
7 Matches 155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
8
9 YZ      825 CATTGGAATCTCTTGGCGCATGTGCCTAAGCAATGGGACATCTTCAGAGAAACTGTT 884
10         11 |||||
11 Db      144 CGAAAGAAATCTTGGCGCATGTGCCTAAGCAATGGGACATCTTCAGAGAAACTGTT 203
12         11 |||||
13 YZ      885 ACTTGATCTCTTAATCCGACAGAGAACATGCGACAGACATCCTGACAGAGAGAGTCA 944
14         11 |||||
15 Db      204 AATGATCTCTTAATCCGACAGAGAACATGCGACAGACATCCTGACAGAGAGAGTCA 263
16         11 |||||
17 YZ      945 GGCGGCAATCTCTTCTCATATCAAAAAGCTCTTCAAGAGAG 982
18         11 |||||
19 Db      264 GGCGGCAATCTCTTCTCATATCAAAAAGCTCTTCAAGAGAGAG 301
20
21 RESULT 3
22 US-09-864-761-24972
23 Sequence 24972 Application US-09864761
24 Patent No. US2002048763A1
25 GenBank Information:
26 APPLICANT: Feng, Sharron G.
27 APPLICANT: Fank, David P.
28 APPLICANT: Huang, David K.
29 APPLICANT: Chen, Wensheng
30 TITLE OF INVENTION: HUMAN GENE EXPRESSION ANALYSIS BY MICROARRAY
31 FILE REFERENCE: Aeomica-X-1
32 CURRENT FILING DATE: 2001-05-23
33 PRIOR APPLICATION NUMBER: US 60/180,312
34 PRIOR FILING DATE: 2000-02-04
35 PRIOR APPLICATION NUMBER: US 60/207,456
36 PRIOR FILING DATE: 2000-05-26
37 PRIOR APPLICATION NUMBER: US 09/632,356
38 PRIOR FILING DATE: 2000-08-03
39 PRIOR APPLICATION NUMBER: CR 24263.6
40 PRIOR FILING DATE: 2000-10-04
41 PRIOR APPLICATION NUMBER: US 00/276,459
42 PRIOR FILING DATE: 2000-09-27
43 PRIOR APPLICATION NUMBER: PCT/US01/00666
44 PRIOR FILING DATE: 2001-01-30
45 PRIOR APPLICATION NUMBER: PCT/US01/00667
46 PRIOR FILING DATE: 2001-01-30
47 PRIOR APPLICATION NUMBER: PCT/US01/00664
48 PRIOR FILING DATE: 2001-01-30
49 PRIOR APPLICATION NUMBER: PCT/US01/00669
50 PRIOR FILING DATE: 2001-01-30
51 PRIOR APPLICATION NUMBER: PCT/US01/00665
52 PRIOR FILING DATE: 2001-01-30
53 PRIOR APPLICATION NUMBER: PCT/US01/00668
54 PRIOR FILING DATE: 2001-01-30
55 PRIOR APPLICATION NUMBER: PCT/US01/00663
56 PRIOR FILING DATE: 2001-01-30
57 PRIOR APPLICATION NUMBER: PCT/US01/00662
58 PRIOR FILING DATE: 2001-01-30
59 PRIOR APPLICATION NUMBER: PCT/US01/00661
60 PRIOR FILING DATE: 2001-01-30
61 PRIOR APPLICATION NUMBER: PCT/US01/00670
62 PRIOR FILING DATE: 2001-01-30
63 PRIOR APPLICATION NUMBER: US 60/234,687
64 PRIOR FILING DATE: 2000-09-21
65 PRIOR APPLICATION NUMBER: US 09/608,408
66 PRIOR FILING DATE: 2000-06-30
67 PRIOR APPLICATION NUMBER: US 09/774,203
68 PRIOR FILING DATE: 2001-01-29
69 NUMBER OF SHO ID NOS: 49117
70 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1

```

```
SEQ ID NO 24072
LENGTH: 152
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008812.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: EST_HUMAN HIT: AV64575.1, EVALU: 2.00e-80
OTHER INFORMATION: SWISSPROT HIT: P26715, EVALU: 6.00e-07
OTHER INFORMATION: NT HIT: AF245219.1, EVALU: 1.00e-80
US-09-864-761-24072
```

```
Query Match 9.94: Score 152; DB 10; Length 152;
Best Local Similarity 100.08; Pred. No. 6,8e-46;
Matches 152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
DB 831 AACGCTGTGGCGGACTGTGCCAAGGACTGGACATCTTCCGAAGAACTGTACTTCA 890
1 AACGCTGTGGCGGACTGTGCCAAGGACTGGACATCTTCCGAAGAACTGTACTTCA 60
DB 891 TCTCTAATCTCCAGTGAATTTGCTATCTCTATCTCTATCTCTATCTCTATCTCTA 960
61 TCTCTAATCTCCAGTGAATTTGCTATCTCTATCTCTATCTCTATCTCTATCTCTA 120
DB 951 ACTGCTGTGAATCAAACTGCTGAGAGGAG 982
121 ACTGCTGTGAATCAAACTGCTGAGAGGAG 152
```

```
RESULT 4
US-09-867-701-7061/c
Sequence 7061, Application US/09467701
Patent No. US2002013227A1
GENERAL INFORMATION:
APPLICANT: Aquale, Paul A.
APPLICANT: Jones, Robert
APPLICANT: Harlocker, Susan L.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 21031.497
CURRENT APPLICATION NUMBER: US/09/867.701
CURRENT FILING DATE: 2001-05-29
NUMBER OF SEQ ID NOS: 10912
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 7061
LENGTH: 472
TYPE: DNA
ORGANISM: Homo sapiens
US-09-867-701-7061
```

```
Query Match 4.28: Score 64; DB 19; Length 472;
Best Local Similarity 64.18; Pred. No. 3.5e-09;
Matches 99; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
```

```
DB 1368 GCACAGGCAATTAATTCCTGCTAATCAATGCTGACAGAGAGGAGCTGAGCTCC 1427
1368 GCACAGGCAATTAATTCCTGCTAATCAATGCTGACAGAGAGGAGCTGAGCTCC 266
DB 1428 ATACAACTCTCTACAGGCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1487
1428 ATACAACTCTCTACAGGCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 266
DB 266 GATGATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 266
1488 TTTGATTTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1524
205 TTTGATTTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 169
```

```
RESULT 5
US-09-905-291A-376
Sequence 376, Application US/09405271A
Patent No. US20020160374A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Holstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gierlsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, Jr.
APPLICANT: Kilaville, Ivan J.
APPLICANT: Maibier, Jennie P.
APPLICANT: Parr, James
APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Thomas, Daniel
APPLICANT: Williams, F. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905.291A
PRIOR APPLICATION NUMBER: P1/US09/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: P2/US09/20594
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: P3/US09/20594
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: P4/US09/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: P5/US09/20594
PRIOR FILING DATE: 1999-09-14
PRIOR APPLICATION NUMBER: P6/US09/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: P7/US09/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: P8/US09/21090
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: P9/US09/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: P10/US09/28413
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: P11/US09/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: P12/US09/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: P13/US09/28565
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: P14/US09/40011
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: P15/US09/40011
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: P16/US09/40011
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: P17/US09/40011
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: P18/US09/40011
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: P19/US09/40011
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 424
SEQ ID NO 376
LENGTH: 997
TYPE: DNA
```


PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/084508
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089908

Query Match 4.28; Score 64.2; EB 12; Length 577;
 Best Local Similarity 52.78; Pred. No. 3.9e-09;
 Matches 188; Conservative 0; Mismatches 163; Indels 7; Gaps 2;

QY 777 AGTCACACAGACAGACAAATCTACAGAGAGCTGACCGATTTGAACACTGCAATTGACGCC 836
 DB 272 AGTTACACTACCTGACAAATTTACAGAGCTCTCTCTACAAATTAAGAACAGGTTACG 331
 QY 837 TGCGCCGACCTGCTCCAGAGCTGAGATTTCTTCAAGTAACTGTTACTATGCTA 896
 DB 332 TCAGCAATTTGTCGATTCAGTACGCGATATTTCAATCAGCTGCTACTCTTTCTA 391
 QY 847 ACTTCACATTTGATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGAT 956
 DB 332 CTGACACATTTCTGCGCGGTTAAGTTAAAGAACTGCTACCGCATGCGGCTACCTGG 451
 QY 957 TCGTATCAAACTGCTGAGAGTACAGAACTGCTACAGCTGACAGCTGCGAGCTAACG 1016
 DB 452 TGGTATCACTACAGAGAGAGAGAGAGATTCCTTCTTACAGAACTGAAATGAGAG 511
 QY 1017 GCTTCTCTGATGAGATTTTCAAGATTTAAATCAAGAGAGAGAGAGAGAGAGAG 1076
 DB 512 AGTTTCTTCTTCTGAG 568
 QY 1077 GCTTACATTTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1133
 DB 569 GCACAGCTTTGACA---AGTCTCTGAGAGCTTTCTGATTTGAGAGAGAGAGAGAGAG 622

RESULT 9

US-09-833-381-995/C
 Sequence 995, Application US/09833381
 Patent No. US20020132090A1
 GENERAL INFORMATION:
 APPLICANT: Robinson, Keith E.
 FILE OF INVENTION: No. US20020132090A1
 CURRENT APPLICATION NUMBER: US/09/833,381
 PRIOR FILING DATE: 2001-04-11
 PRIOR APPLICATION NUMBER: 09/516,448
 NUMBER OF SEQ ID NOS: 2050
 SOFTWARE: FASTSEQ for Windows Version 3.0
 SEQ ID NO 995
 LENGTH: 528
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1) ... (528)
 OTHER INFORMATION: n = A,T,C or G
 US-09-833-381-995

Query Match 3.68; Score 56; DB 10; Length 528;
 Best Local Similarity 52.58; Pred. No. 7.5e-07;
 Matches 147; Conservative 0; Mismatches 130; Indels 3; Gaps 1;

QY 807 TGACGATTTAAGACATTTGATTAAGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 866
 DB 355 TCAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 236
 QY 867 TCTTCAGAGAACTGTTACTTCAATGCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 926
 DB 295 GATTCGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCT 236
 QY 927 GCGGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCT 966

DB 235 AGACCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 331
 QY 987 TCTTCAGAGAACTGTTACTTCAATGCTACAGAGAGAGAGAGAGAGAGAGAGAGAG 1076
 DB 175 TCATACAG 391
 QY 1047 ACTTCACATTTGATTTGATTAATTTGATTAATTTGATTAATTTGATTAATTTGAT 956
 DB 118 AAGCTAATGAGAAATGCGAGATGATGATGATGATGATGATGATGATGATGATGAT 79

RESULT 10

US-09-880-107-3731
 Sequence 3731, Application US/09880107
 Patent No. US20020142981A1
 GENERAL INFORMATION:
 APPLICANT: Horne, Darci T.
 APPLICANT: Vockley, Joseph G.
 APPLICANT: Scherf, Uwe
 TITLE OF INVENTION: Gene Expression Profiles in Liver Can-
 CURRENT APPLICATION NUMBER: US/09/880,107
 PRIOR FILING DATE: 2001-06-14
 PRIOR APPLICATION NUMBER: US 60/211,379
 PRIOR FILING DATE: 2000-06-14
 PRIOR APPLICATION NUMBER: US 60/217,054
 PRIOR FILING DATE: 2000-10-02
 NUMBER OF SEQ ID NOS: 3950
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 3731
 LENGTH: 1300
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: Genbank Accession No. US20020142981A1
 US-09-880-107-3731

Query Match 3.44; Score 52; DB 10; Length 1
 Best Local Similarity 58.34; Pred. No. 2.1e-05;
 Matches 91; Conservative 0; Mismatches 65; Indels 0;

QY 847 CTGCTCCAG 331
 DB 646 CTGCTCCAG 331
 QY 907 GACTGCTACAG 331
 DB 706 GCTGCTGCTGAG 331
 QY 967 AAGCTAATGAGAAATGCGAGATGATGATGATGATGATGATGATGATGATGATGAT 79
 DB 766 CTGCTGCTGAG 331

RESULT 11

US-09-880-107-2235
 Sequence 2235, Application US/09880107
 Patent No. US20020142981A1
 GENERAL INFORMATION:
 APPLICANT: Horne, Darci T.
 APPLICANT: Vockley, Joseph G.
 APPLICANT: Scherf, Uwe
 TITLE OF INVENTION: Gene Expression Profiles in Liver Can-
 CURRENT APPLICATION NUMBER: US/09/880,107
 PRIOR FILING DATE: 2001-06-14
 PRIOR APPLICATION NUMBER: US 60/211,379
 PRIOR FILING DATE: 2000-06-14
 PRIOR APPLICATION NUMBER: US 60/217,054
 PRIOR FILING DATE: 2000-10-02

```

query Match      3 38,      Score= 51.2,  PR 10;  Length 3750;
Heat Local Similarity 44.1%;  Pred. No. 6.9e-05;
Matches 261; Conservative 0; Mismatches 328; Indels 3; Gaps 1;

```

	Query Match	4.3%	Score 51	DB 105	Length 443
	Best Local Similarity	58.1%	Fred No. 2.1e-05		
	Matches	967	Conservative 0	Mismatches 65	Gaps 0
QY	847	CCTGTCACAGCATGTGCATTTTCCAGGAATGGTACTCATTGCTACACGCCAGCG	906		
IB	77	CTCCCAACAAACAGCAGCAATACCAAGCCAGCGTGCTACGCTCTGCTGCGCGGAAA	136		
QY	907	GAACTGATTAATCTGATATATGCTGTTAATGATGAGAGAGGACATCGTGATATCAA	966		

```

db 137 GGCCTGGCCAGAGGCTGACAGACTGCTGAGTGGAGAAATGCTTGTGTGTGG 196
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 967 AACTGCTAGAGAGCAAACTTCTTAACTTGTGAA 1001
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
db 197 CTCCTGGAGAGAGCAAAATTTATCCAGCACCA 231

```

RESULT 14

```

US-09-862-802-9
: Sequence 9, Application US/09862802
: Patent No. US20020165346A1
: GENERAL INFORMATION:
: APPLICANT: Schering-Plough Corporation
: TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
: FILE REFERENCE: SF0695B
: CURRENT APPLICATION NUMBER: US/09/862,802
: CURRENT FILING DATE: 2002-01-10
: PRIOR APPLICATION NUMBER: US 09/111,470
: PRIOR FILING DATE: 1998-07-08
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 9
: LENGTH: 1370
: TYPE: DNA
: ORGANISM: unknown
: FEATURE:
: OTHER INFORMATION: mammalian nucleic acid
: NAME/KEY: CDS
: LOCATION: 273...1091
: OTHER INFORMATION: protein coding sequence
US-09-862-802-9

```

```

Query Match 3.38; Score 50.4; DB 9; Length 1370;
Best Local Similarity 58.88; Pred. No. 6.4e-05;
Matches 87; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

```

```

QY 847 CTGTCCAGAGACTGACATCTTCGACGAAAGTGTACTGATGTCTAACGCCACG 906
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
db 683 CTCCTGGAGAGAGTGTGAGAGCAAGACAGCTGTACTGTTCTGACAGCTGCGAT 742
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 907 GAATGGCAGACATCTGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
db 743 GTCTGTGGCGAGAGCTGACAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 802
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 967 AACTGTAGAGAGAGCAAACTTCTTAACTG 994
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
db 803 CTCAGGAGAGAGCAAAATTTGTCCAG 830

```

RESULT 15

```

US-09-862-802-3
: Sequence 3, Application US/09862802
: Patent No. US20020165346A1
: GENERAL INFORMATION:
: APPLICANT: Schering-Plough Corporation
: TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
: FILE REFERENCE: SF0695B
: CURRENT APPLICATION NUMBER: US/09/862,802
: CURRENT FILING DATE: 2002-01-10
: PRIOR APPLICATION NUMBER: US 09/111,470
: PRIOR FILING DATE: 1998-07-08
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3
: LENGTH: 1458
: TYPE: DNA
: ORGANISM: unknown
: FEATURE:
: OTHER INFORMATION: mammalian nucleic acid
: NAME/KEY: CDS
: LOCATION: 257...1204
: OTHER INFORMATION: protein coding sequence

```

```

: NAME/KEY: variation
: LOCATION: 608...673
: OTHER INFORMATION: short form lacks these nucleotides
US-09-862-802-3

```

```

Query Match 3.38; Score 50.4; DB 9; Length
Best Local Similarity 58.88; Pred. No. 6.4e-05;
Matches 87; Conservative 0; Mismatches 61; Indels

```

```

QY 847 CTGTCCAGAGACTGACATCTTCGACGAAAGTGTACTGATGTCTAACGCCACG 906
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
db 796 CTCCTGGAGAGAGTGTGAGAGCAAGACAGCTGTACTGTTCTGACAGCTGCGAT 742
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 907 GAATGGCAGACATCTGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
db 856 GTCTGTGGCGAGAGCTGACAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 802
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 967 AACTGTAGAGAGAGCAAACTTCTTAACTG 994
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
db 916 CTCAGGAGAGAGCAAAATTTGTCCAG 830

```

```

Search completed: December 7, 2002, 06:10:15
Job time : 78 secs

```

The first part of the paper discusses the importance of the study and the objectives of the research. It then proceeds to a literature review, followed by a description of the methodology used in the study. The results of the study are then presented, followed by a discussion of the findings and their implications. Finally, the paper concludes with a summary of the main points and a list of references.

The study was conducted in a laboratory setting, and the results were compared with those of previous studies. The findings of the study are consistent with those of previous studies, and they provide new insights into the phenomenon being studied. The implications of the study are discussed in detail, and the authors conclude that the study has made a significant contribution to the field.

The authors would like to thank the following people for their assistance in the study: [Name], [Name], and [Name]. They would also like to thank the following organizations for their support: [Organization], [Organization], and [Organization].

Query Match	Best Local Similarity	88.8%	Score 665.4	DP 4	Length 1212			
Query Match	Best Local Similarity	88.8%	Score 665.4	DP 4	Length 1212			
Matches 720	Conservative	0	Mismatches	91	Indels	0	Gaps	0
QY 424	TCACGACATATCCGACACACGCCAATCTACACAGAACTGACGCCACCGTTAAAGCTGACT	483						
DB 342	TCACGACAAATCTCAATCTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT	401						
QY 484	GCGTCAGCTCTCAGACAAATCTCAATCAATCAATCAATCAATCAATCAATCAATCAAT	543						
DB 402	GCGTCAGCTCTCAGACAAATCTCAATCAATCAATCAATCAATCAATCAATCAATCAAT	441						
QY 544	GATTCATATAGGACTAT	603						
DB 462	GATTCATATAGGACTAT	521						
QY 604	CGGCTGACAGGCTGACAGGCTGACAGGCTGACAGGCTGACAGGCTGACAGGCTGACAG	663						
DB 522	CGGCTGACAGGCTGACAGGCTGACAGGCTGACAGGCTGACAGGCTGACAGGCTGACAG	561						
QY 664	GATTCATATAGGACTAT	723						
DB 582	GATTCATATAGGACTAT	641						
QY 724	GATTCATATAGGACTAT	783						
DB 642	GATTCATATAGGACTAT	701						
QY 784	GATTCATATAGGACTAT	843						
DB 702	GATTCATATAGGACTAT	761						
QY 844	GATTCATATAGGACTAT	903						
DB 762	GATTCATATAGGACTAT	821						
QY 904	GATTCATATAGGACTAT	963						
DB 822	GATTCATATAGGACTAT	861						
QY 964	GATTCATATAGGACTAT	1023						
DB 882	GATTCATATAGGACTAT	941						
QY 1024	GATTCATATAGGACTAT	1083						
DB 942	GATTCATATAGGACTAT	1001						
QY 1084	GATTCATATAGGACTAT	1143						
DB 1062	GATTCATATAGGACTAT	1061						
QY 1144	GATTCATATAGGACTAT	1203						
DB 1062	GATTCATATAGGACTAT	1121						
QY 1204	GATTCATATAGGACTAT	1263						
DB 1122	GATTCATATAGGACTAT	1152						

RESULT 4

US-09-517-605-1

Sequence 1, Application US/09517605

Patent No. 6391567

GENERAL INFORMATION:

APPLICANT: Liltman, Dan R.

APPLICANT: Kwon, Douglas S.

APPLICANT: van Kooyk, Yvette

APPLICANT: Gelfenbeck, Tracy

TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO

TITLE OF INVENTION: CELLS

FILE REFERENCE: 1049-1-017

CURRENT APPLICATION NUMBER: US/09/517,605

CURRENT FILING DATE: 2000-04-02

NUMBER OF SEQ ID NOS: 17

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 1

LENGTH: 1312

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (42)..(1253)

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

US-09-517-605-1

RESULT 4

US-09-591-435-11

Sequence 11, Application US/09591435

Patent No. 6280953

GENERAL INFORMATION:

APPLICANT: MESSIER, WALTER

APPLICANT: STRELA, JAMES M

TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AN

TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED W

FILE REFERENCE: GPO 200 2

CURRENT APPLICATION NUMBER: US/09/591,435

CURRENT FILING DATE: 2000-06-09

PRIOR APPLICATION NUMBER: 09/591,435

PRIOR FILING DATE: 2000-06-09

PRIOR APPLICATION NUMBER: 09/240,915

PRIOR FILING DATE: 1999-01-29

PRIOR APPLICATION NUMBER: 60/074,987

PRIOR FILING DATE: 1998-01-30

PRIOR APPLICATION NUMBER: 60/098,987

PRIOR FILING DATE: 1998-09-02

NUMBER OF SEQ ID NOS: 13

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 11

LENGTH: 1212

TYPE: DNA

ORGANISM: Gallus gallus

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

US-09-591-435-11

[illegible]

1 RESULT 6,
 2 US-08-772-449-20
 3 Sequence 20, Application 35/28772440
 4 Patent No. 6046158
 5
 6 GENERAL INFORMATION:
 7
 8 APPLICANT: Aritzumi, Kiyoshi
 9
 10 APPLICANT: Takashima, Akira
 11
 12 TITLE OF INVENTION: UNIQUE DENDRITIC CELL-ASSOCIATED C-TYPE
 13
 14 TITLE OF INVENTION: LECTIN, DECTIN-1 AND DECTIN-2 COMPOSITIONS AND USES
 15
 16 NUMBER OF SEQUENCES: 42
 17
 18 CORRESPONDENCE ADDRESS:
 19
 20 ATTORNEY: Aritzumi, Kiyoshi & Fuchise
 21
 22 STREET: P.O. Box 4434
 23
 24 CITY: Houston
 25
 26 STATE: Texas
 27
 28 COUNTRY: USA
 29
 30 ZIP: 77210
 31
 32 COMPUTER READABLE FORM:
 33
 34 MEDIUM TYPE: Floppy disk
 35
 36 COMPUTER: IBM PC compatible
 37
 38 OPERATING SYSTEM: PC-DOS/MS-DOS
 39
 40 SOFTWARE: PatentIn Release #1.0, Version #1.30
 41
 42 CURRENT APPLICATION DATA:
 43
 44 APPLICATION NUMBER: US-2007/772,440
 45
 46 FILING DATE: CONCURRENTLY HERewith

TELEFAX: 212 818-9479
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1897 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEetical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Bos taurus
: TISSUE TYPE: Vascular endothelial cells
: IMMEDIATE SOURCE:
: LIBRARY: Bovine aortic endothelial cell cDNA
: CLONE: pBlox-1
: FEATURE:
: NAME/KEY: polyA-site
: LOCATION: 1880..1897
: FEATURE:
: NAME/KEY: misc_PNA
: LOCATION: 1859..1864
: OTHER INFORMATION: /function: "PolyA signal"
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..34
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 848..1897
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 35..847
: US-08-809-494A-1

Query Match
Best local similarity 48.88; Pred No 3.5e-05;
Matches 147; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
3.58; Score 54.6; DB 2; Length 1897;

QY 798 ACCGAGGACTGACGATTGAAAGTTCATTTAAAGGCTGAGGAGTGGAGTGGAGGAGG 857
DB 402 ACTGAAATCTGCAAGAGTCTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 461
QY 858 ACTGAGACTCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 917
DB 462 ACTGGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 521
QY 918 ACTGGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 977
DB 522 AAACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 581
QY 978 ACTGAGACTCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1037
DB 582 AACTGGAATTCATGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 641
QY 1038 GATGAGCTTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1097
DB 642 CAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 701
QY 1098 T 1098
DB 702 T 702

CORRESPONDENCE ADDRESS:
: ADDRESSER: McAdams Fisher Nissen Goldberg & Kiel
: STREET: 261 Madison Avenue
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10016-2391
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 08/073522,302
: FILING DATE: 12-JUL-1999
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 6-421705
: FILING DATE: 30-NOV-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 7-214206
: FILING DATE: 31-JUL-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Goldberg, Jules E.
: REGISTRATION NUMBER: 24408
: REFERENCE/DOCKET NUMBER: JG-YY-4364PCT/D
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212 986-4090
: TELEFAX: 212 818-9479
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1897 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEtical: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Bos taurus
: TISSUE TYPE: Vascular endothelial cells
: IMMEDIATE SOURCE:
: LIBRARY: Bovine aortic endothelial cell cDNA
: CLONE: pBlox-1
: FEATURE:
: NAME/KEY: polyA-site
: LOCATION: 1880..1897
: FEATURE:
: NAME/KEY: misc_PNA
: LOCATION: 1859..1864
: OTHER INFORMATION: /function: "PolyA signal"
: FEATURE:
: NAME/KEY: 5'UTR
: LOCATION: 1..34
: FEATURE:
: NAME/KEY: 3'UTR
: LOCATION: 848..1897
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 35..847
: US-09-352-302-1

Query Match
Best local similarity 48.88; Pred. No. 3.5e-05;
Matches 147; Conservative 0; Mismatches 154; Indels 0; Gaps 0;
3.58; Score 54.6; DB 4; Length 1897;

QY 798 ACCGAGGACTGACGATTGAAAGTTCATTTAAAGGCTGAGGAGTGGAGTGGAGGAGG 857
DB 402 ACTGAAATCTGCAAGAGTCTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 461
QY 858 ACTGAGACTCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 917
DB 462 ACTGGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 521


```

APPLICATION NUMBER: JP 6-321705
FILING DATE: 30-NOV-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 7 214206
FILING DATE: 31-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldberg, Jules E
REGISTRATION NUMBER: 24408
REFERENCE/DOCKET NUMBER: JG-YY-4363PCT/D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 986-4090
TELEFAX: 212 818-9479
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1906 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bos taurus
TISSUE TYPE: Vascular endothelial cells
IMMEDIATE SOURCE:
LIBRARY: Bovine aortic endothelial cells cDNA
CLONE: pBlox-1
FEATURE:
NAME/KEY: polyA_site
LOCATION: 1889..1906
FEATURE:
NAME/KEY: misc_PNA
LOCATION: 1864..1873
OTHER INFORMATION: /function= "Polya signal"
NAME/KEY: 5'UTR
LOCATION: 1..34
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 857..1906
FEATURE:
NAME/KEY: CDS
LOCATION: 35..856
US-09-352-302-4

```

```

Query Match: 3.5%, Score 54.6, DB 4, Length 1906,
Best Local Similarity 48.8%, Pred. No. 3 50-05;
Matches 147; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

```

```

798 ACCGCTACCTGACCGATTTTAACTATTTGACCGCTCTGCGGCGCACTTCTTCAACG 257
DB 411 ACCTGATCTCTCAAGAGTTTGTAAAGAGGCAACAACTATTCAAGTCTGCGGCAAG 476
QY 858 ACTGACATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 917
DB 471 ACTGATCTCTCAAGAGTATTATGATGTTAACTTCTGATGAGTCTTATGATGAAA 500
QY 918 ACTGATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 977
DB 531 AAAGGACAGTACGAGTCTCTTCTTCTGATGTTAACTTCAAGAGGAGTCTGATG 590
QY 978 AATGACAGTCTCTCAAGTCTGATGTTAACTTCAAGAGGAGTCTGATG 1047
DB 591 AATGACAGTCTCTCAAGTCTGATGTTAACTTCAAGAGGAGTCTGATG 650
QY 1038 CAGACGATTAATCAAGAGTCTGATGTTAACTTCAAGAGGAGTCTGATG 1047
DB 651 CAGACGATTAATCAAGTCTGATGTTAACTTCAAGAGGAGTCTGATG 710
QY 1098 T 1098
DB 711 T 711

```

```

RESULT 12
US-08-365-1038-9
Sequence 3, 64111 at: 00 55/062651038
Patent No. 5766944
GENERAL INFORMATION:
APPLICANT: Lynch, Richard G.
APPLICANT: Lynch, Raphael D.
APPLICANT: Yodanis, Jundi
TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
TITLE OF INVENTION: and Methods of Use for Same
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSER: Zelig, Mickey, Thomas, Voorhees & Sease
STREET: 801 Grand Ave., Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: P-Pos/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/365,1038
FILING DATE: 28-DEC-1994
CLASSIFICATION: A35
ATTORNEY/AGENT INFORMATION:
NAME: Nobel, Heidi S.
REGISTRATION NUMBER: 37,719
PREFERENCE/DOCKET NUMBER: 0117 N5-24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 288-3667
TELEFAX: (515) 288-1338
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1025 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 45..1025
US-08-365-1038-9

```

```

Query Match: 3.5%, Score 54.2, DB 1, Length 1025;
Best Local Similarity 51.9%, Pred. No. 3 50-05;
Matches 150; Conservative 0; Mismatches 133; Indels 0; Gaps 1;

```

```

408 GACGATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 567
DB 508 GACGATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 568
QY 568 GACGATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 927
DB 569 GACGATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 928
QY 928 GACGATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 987
DB 629 GACGATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 988
QY 988 GACGATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 1047
DB 629 GACGATCTCTCAAGAGTATTATGATGTTAACTTCAAGAGGAGTCTGATG 742
QY 1048 CAGACGATTAATCAAGAGTCTGATGTTAACTTCAAGAGGAGTCTGATG 1096
DB 743 CAGACGATTAATCAAGAGTCTGATGTTAACTTCAAGAGGAGTCTGATG 791

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: December 7, 2002, 06:07:57 : Search time 2956 Seconds

(without alignments)
15181 504 Million cell updates/sec

Title:	US-09-831-458A-25
Portfolio cover:	1543

Sequence:

[illegible]

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Number of hits satisfying chosen parameters	Number of hits satisfying chosen parameters
4104280	4104280

Minimum DB seq length: 0

Maximum I/O seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

1)atäbäse:

1:	qb_ba:*
2:	qb_htg:*
3:	qb_in:*
4:	qb_om:*
5:	qb_ov:*
6:	qb_pat:*
7:	qb_ph:*
8:	qb_pl:*
9:	qb_pr:*
10:	qb_ro:*
11:	qb_sts:*
12:	qb_sy:*
13:	qb_un:*
14:	qb_vl:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_ph:*
24:	em_pl:*
25:	em_ro:*
26:	em_sts:*
27:	em_sy:*
28:	em_un:*
29:	em_vl:*
30:	em_htg_hum:*
31:	em_htg_in:*
32:	em_htg_other:*
33:	em_htg_mu:*
34:	em_htg_pl:*
35:	em_htg_ro:*
36:	em_htg_sts:*
37:	em_htg_vl:*
38:	em_sy:*
39:	em_htg_hum:*
40:	em_htg_in:*
41:	em_htg_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the rest, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Query Match	Length	DBs	DB	
1	1054.8	68.4	1510	%	AF-z45819	A
2	907.2	58.8	1860	%	AF-z90887	A
3	831.2	53.9	999	%	AY042238	A
4	780.2	50.6	1290	%	AY042234	A
5	724.4	47.0	1105	%	AY042231	A
6	716.2	46.4	1643	%	AX-z57118	A
7	716.2	46.4	1643	%	AB015629	A
8	668.4	43.3	1146	%	AF-z69755	A
9	667.6	43.3	1166	%	AY040319	A
10	665.4	43.2	1083	%	AY042236	A
11	665.4	43.2	1113	%	AY042235	A
12	665.4	43.2	1212	%	AK146105	A
13	665.4	43.2	1212	%	AX-z42003	A
14	665.4	43.2	1215	%	AX0199403	A
15	665.4	43.2	1215	%	AY042231	A
16	665.4	43.2	1234	%	AY-z42230	A
17	665.4	43.2	1312	%	AX-z87103	A
18	665.4	43.2	1312	%	HDMLECT1N0	A
19	665.4	43.2	1315	%	AY-z42239	A
20	665.4	43.2	4266	%	AF-z90886	A
21	662.8	43.0	1146	%	AF-z42372	A
22	659	42.7	1212	%	AK146107	A
23	659	42.7	1212	%	AX-z42005	A
24	659	42.7	1212	%	AF-z91085	A
25	653.8	42.4	1212	%	AK146106	A
26	653.8	42.4	1212	%	AX-z42004	A
27	652.6	42.3	1215	%	AF-z91086	A
28	619.4	40.2	1197	%	AY042232	A
29	564	36.6	792	%	AY042230	A
30	547.4	35.6	807	%	AY042237	A
31	545.8	35.4	939	%	AY042223	A
32	497.8	32.3	1535	%	AY042236	A
33	497.8	32.3	1648	%	AY042235	A
34	464.6	30.1	994	%	AY042232	A
35	420	27.2	1472	%	AY042239	A
36	420	27.2	4506	%	AF-z90886	A
37	420	27.2	143619	%	AF-z90887	A
38	415.2	26.9	190628	%	AK-z42700	A
39	412.4	26.7	2991	%	AK-z53750	A
40	409.4	26.5	811	%	AY-z4237	A
41	389.6	25.3	143619	%	AC000812	A
42	389.6	25.3	17792	%	AF-z908873	A
43	388.2	25.2	4434	%	AF-z909479	A
44	303	19.6	587	%	AF-z42235	A
45	257.4	16.7	927	%	AF-z74471	A

APPENDIX A

RESULT 1	AF245219	1510 bp	mRNA	1 times	JAN 2001
LOCUS	AF245219				NC mRNA
DEFINITION	AF245219 Homo sapiens probable mucinose-binding C-type 10 complete cds.				
ACCESSION	AF245219				
VERSION	AF245219.2	GI:12081796			
KEYWORDS					
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea (1 bases 1 to 1510)				
REFERENCE	Soilleux,E.J., Batten,K. and Frowsdale,J.				
AUTHORS	DC-STON; a related gene, DC-STONR, and CD24 for				
TITLE					

[illegible]

[illegible][illegible]

/chromosome="19"
 /map="19p13.3"
 1..1215
 /gene="CD209"
 1..1215
 /gene="CD209"
 /note="similar to the nucleotide sequence deposited in
 GenBank Accession Number AF008412. alternatively spliced"
 /rnodon.start=1
 /evidence=experimental
 /product="mdc-SIGNIA type I isoform"
 /protein_id="AAK91846.1"
 /db_xref="GI:15281073"
 /translation="MDSKRPRLQGLHLEEDLGLSHKQIPKYSIAQLGHPYV
 LQLSFTLLAGLIVYKVPSSISOESRODAIYONLTOLKAAVGELESEKIOEIV
 ETLKAAVGELEPEKSKIOEIVLRLKAAVGELEPEKSKIOEIVLRLKAAVGELE
 PEKSKIOEIVLRLKAAVGELEPEKSKIOEIVLRLKAAVGELEPEKSKIOEIV
 LRLKAAVGELEPEKSKIOEIVLRLKAAVGELEPEKSKIOEIVLRLKAAVGELE
 WHDSITACKVQQLVYIKSAEONFLQLOSRSNRPFTWGLSDIHDGCTWQWDGSP
 LIPSPKQIMKQGNVNGEELDAEFSQNGMNDKCNIAKFIKKSAASCSKDEEPL
 SPAAPTPNPPEPA"

BASE COUNT 315 a 321 c 349 g 230 t
 ORIGIN

Query Match 43.2%; Score 665.4; DB 9; Length 1215;
 Post-local Similarity 88.8%; Pred. No. 1,1e-162;
 Matches 720; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 424 TCAGCAATCGAGTAAATCTAATCTAGAACTGAATCTGAATTAATCTGAACT 483
 DB 342 TCAGAGAAATCTAGCTGCAGAGATCTACTAGTAACTGATCTGAAAGCTGCAGT 401
 QY 484 GGGTAGCTTCAGAGAAATCCAGCTGCAGAGATCTACAGAGCTGAGCCAGCTGAA 543
 DB 402 GGGTAGCTTCAGAGAAATCTAAGCTGCAGAGATCTAAGAGCTGAGCTGAA 461
 QY 544 GGGTAGCTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAG 603
 DB 462 GGGTAGCTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAG 521
 QY 604 GGGTAGCTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAG 663
 DB 522 TCGGCTGAAGCTGCAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAG 581
 QY 664 GGGTAGCTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAG 723
 DB 582 GGGTAGCTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAGTGGTGAAG 641
 QY 724 GATCTACAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 783
 DB 642 GATCTACAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 701
 QY 784 GAGTAGTAATCTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 843
 DB 702 GAGTAGTAATCTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 761
 QY 844 GAGTAGTAATCTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 903
 DB 762 GAGTAGTAATCTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 821
 QY 904 GAGTAGTAATCTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 963
 DB 822 GAGTAGTAATCTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 881
 QY 964 GAGTAGTAATCTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1024
 DB 882 GAGTAGTAATCTAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 941
 QY 1024 CTGATGGAGATTTGATGATTTAAATAGAGAGCTAGAGAGCTAGAGAGCTAGAG 1083
 DB 942 CTGATGGAGATTTGATGATTTAAATAGAGAGCTAGAGAGCTAGAGAGCTAGAG 1001
 QY 1084 TCTGTACCCAGCTTCAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 1143

DB 1002 TCTGT 1061
 QY 1144 AGAGT 1203
 DB 1062 AGAGT 1121
 QY 1204 CTGATGT 1263
 DB 1122 CTGATGT 1181

Search completed: December 7, 2002, 07:57:54
 Job time : 2976 secs

...

Tue Dec 10 10:10:09 2002

us-09-831-458a-12.std.rapb

Page 1

GenCore version 5.1.3
Copyright (c) 1993-2002 Compugen Ltd

OM protein - protein search, using sw model

Run on: December 7, 2002, 11:02:30 Search time: 21 seconds

(without alignments)
251,369 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 1696
Sequence: 1 MSDSKEPRVQDLGLGCLSH PCDVNYWICKKPKACFCE 325

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

Published Applications-AA:
1. /cgn2_6/ptodata/2/pubpaa/US08_NEW_PRR pep: *
2. /cgn2_6/ptodata/2/pubpaa/PT_NEW_PRR pep: *
3. /cgn2_6/ptodata/2/pubpaa/US06_NEW_PRR pep: *
4. /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep: *
5. /cgn2_6/ptodata/2/pubpaa/US07_NEW_PRR pep: *
6. /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep: *
7. /cgn2_6/ptodata/2/pubpaa/PTGT_PUBCOMB pep: *
8. /cgn2_6/ptodata/2/pubpaa/US08_NEW_PRR pep: *
9. /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep: *
10. /cgn2_6/ptodata/2/pubpaa/US10_NEW_PRR pep: *
11. /cgn2_6/ptodata/2/pubpaa/US10_PRR pep: *
12. /cgn2_6/ptodata/2/pubpaa/US10_PRR pep: *
13. /cgn2_6/ptodata/2/pubpaa/US60_NEW_PRR pep: *
14. /cgn2_6/ptodata/2/pubpaa/US60_PRR pep: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	319.5	18.8	316	9	US-09-862-802-4 Sequence 4, Appl1
2	305	18.0	273	9	US-09-862-802-10 Sequence 10, Appl1
3	289.5	17.1	293	9	US-09-978-295A-231 Sequence 231, App
4	289.5	17.1	293	9	US-09-978-297-231 Sequence 231, App
5	289.5	17.1	293	9	US-09-978-297-231 Sequence 231, App
6	284	16.7	50	10	US-09-864-701-40945 Sequence 40945, A
7	274.5	16.2	333	10	US-09-764-870-279 Sequence 279, App
8	273.5	16.1	219	9	US-09-905-291A-377 Sequence 377, App
9	273.5	16.1	219	10	US-09-904-420-377 Sequence 377, App
10	273.5	16.1	219	10	US-09-909-088B-377 Sequence 377, App
11	273.5	16.0	291	9	US-09-862-802-24 Sequence 24, Appl1
12	271.5	15.8	287	9	US-09-862-802-5 Sequence 5, Appl1
13	267.5	15.5	182	12	US-10-040-466-4 Sequence 4, Appl1
14	263.5	15.5	213	12	US-10-040-466-2 Sequence 2, Appl1
15	263.5	15.5	213	12	US-10-040-466-2 Sequence 2, Appl1
16	246	14.5	273	10	US-09-79b-858-47 Sequence 47, Appl1
17	234.5	13.8	493	9	US-09-870-759-142 Sequence 142, Appl1
18	232	13.7	237	9	US-09-862-802-2 Sequence 2, Appl1
19	232	13.7	237	9	US-09-870-759-49 Sequence 49, Appl1

20	232	13.7	243	10	US-09-764-870-464 Sequence 464, Appl1
21	232	13.7	243	10	US-09-764-870-303 Sequence 303, Appl1
22	230	13.6	457	10	US-09-745-763-218 Sequence 218, Appl1
23	230	13.6	542	10	US-09-745-763-219 Sequence 219, Appl1
24	226.5	13.4	284	9	US-09-862-802-8 Sequence 8, Appl1
25	210.5	12.4	1456	9	US-09-870-755-95 Sequence 95, Appl1
26	206	12.1	149	10	US-09-944-807-6 Sequence 6, Appl1
27	202.5	11.9	154	10	US-09-764-870-439 Sequence 439, Appl1
28	202.5	11.9	154	10	US-09-764-870-439 Sequence 439, Appl1
29	189	11.1	198	10	US-09-764-870-439 Sequence 439, Appl1
30	189	11.1	203	10	US-09-764-870-439 Sequence 439, Appl1
31	188.5	11.1	232	12	US-10-052-586-94 Sequence 94, Appl1
32	186	11.0	182	10	US-09-764-870-410 Sequence 410, Appl1
33	186	11.0	189	9	US-09-811-367B-1 Sequence 1, Appl1
34	183.5	10.8	271	9	US-09-922-598-357 Sequence 357, Appl1
35	183.5	10.8	271	9	US-09-989-293A-357 Sequence 357, Appl1
36	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1
37	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1
38	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1
39	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1
40	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1
41	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1
42	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1
43	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1
44	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1
45	183.5	10.8	271	10	US-09-989-723-357 Sequence 357, Appl1

ALIGNMENTS

RESULT 1
US-09-862-802-4
Sequence 4, Application US-09862802
Patent No. US2002065346A1
GENERAL INFORMATION:
APPLICANT: Schering-Plough Corporation
TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES, REI
FILE REFERENCE: SP06958
CURRENT APPLICATION NUMBER: US-09/862,802
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US-09/411,470
PRIOR FILING DATE: 1998-07-08
NUMREP OF SEQ ID NOS: 11
SOFTWARE: Patent version 3.1
SEQ ID NO: 4
LENGTH: 316
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: mammalian protein
US-09-862-802-4

Query Match	18.8%	Score 319.5	DB 9	Length
Best Local Similarity	28.8%	Prod. No. 34e-18		
Matches	90	Conservative	53	Mismatches 103
DB	9	US-09-862-802-4		
Q	23	IVIGLISPMIAGVVALIVVSKVPSISQSEEDATYONLTOLKAV		
Y	23	IVIGLISPMIAGVVALIVVSKVPSISQSEEDATYONLTOLKAV		
DB	41	ILISTGLILVILVGVGHSRKFQDVLTRD-----FSNFTSTVA		
Q	83	ELVGLDGLKAAVGLLEFESKTFYETLTKAAVGLPEKSKLQIV		
Y	83	ELVGLDGLKAAVGLLEFESKTFYETLTKAAVGLPEKSKLQIV		
DB	89	ALISGSSILGLVAD--FAVVDGRG--KAGVSEI-----		
Q	143	ELPE-----KKEIYELGLTKAAVGLPPGSKQGLYQI		
Y	143	ELPE-----KKEIYELGLTKAAVGLPPGSKQGLYQI		
DB	134	HCHGCSVCVPAHSEMLRVOGLVQDLKLLTGATLNNNA-----		
Q	191	CPHCPKMTPTPCNCLYKMSNSQRMHSVTCVQVQGLVITADPR		
Y	191	CPHCPKMTPTPCNCLYKMSNSQRMHSVTCVQVQGLVITADPR		
DB	180	CTCTVWVHDSCTVPHSCKMAVAFKVCGLKMLVIVNSRPPQN		

```

QY 251 FSWKGLSDLNQECTWQWQVDSPLSPFCQKYNSSGPN-- SGNEDCAFSSGSGWR 103
      ++++++ ++++++ ++++++ ++++++ ++++++
DB 236 YIMWGLST--PTAKMKVVFVAVATSPQN-WKEDQPTWQWQVQSGHVSQDRCAPHPVQPMN 292
      ++++++ ++++++ ++++++ ++++++ ++++++
QY 304 DNRQDVINWICK 316
      ++++++ ++++++
DB 293 DDVQGRFYHWCE 305

RESULT 2
US-09-862 802 10
: Sequence 10, Application US/09862802
: Patent No. US20020165346A1
: GENERAL INFORMATION:
: APPLICANT: Schering-Plough Corporation
: TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES: PEPTIDE PEPTIDES
: FILE REFERENCE: SF0695B
: CURRENT APPLICATION NUMBER: US/09/862,802
: PRIOR FILING DATE: 2002-01-10
: PRIOR APPLICATION NUMBER: US 09/111,470
: PRIOR FILING DATE: 1998-07-08
: NUMBER OF SEQ. ID NOS: 11
: SOFTWARE: PatentIn version 3.1
: SEQ. ID NO: 10
: LENGTH: 273
: TYPE: PRT
: ORGANISM: monkey
: FEATURE:
: OTHER INFORMATION: mammalian protein
US-09-862-802 10

Query Match 18.08, Score 305, DB 9, Length 273,
Best local similarity 29.38, Pctd. No. 3, gc-17,
Matches 89; Conservative 49; Mismatches 96; Indels 70; Gaps 15:

QY 20 HGAIVGLLTFMLAGVAILVQVSKVPSISQBSQEDALTYQNTQKAAGLSLKS 75
      ++++++ ++++++ ++++++ ++++++ ++++++
DB 22 NGPLPLOS--LIVYIVGVGRQNSKFORDLTLD--FSNT-----SN 63
      ++++++ ++++++ ++++++ ++++++ ++++++
QY 80 KLOIYQELTQKAAGELPEKSKIOEYQELTRKAAGELPFKSLQEIYQELPLKA 139
      ++++++ ++++++ ++++++ ++++++ ++++++
DB 64 TVARI-QALIS-----QTSIPE--TASIKAEVSGF--KQFRQAVHSPM----- 104
      ++++++ ++++++ ++++++ ++++++ ++++++
QY 140 AVATPFPSKSLQEIYQELTQKAAGVLEPQSKQQTQYPTIDTKATPELQPHCYQMT 199
      ++++++ ++++++ ++++++ ++++++ ++++++
DB 105 -----LIVQQLVQDQKLKLTQVAVTLNNNEE-----ASTEGTC--QPVNV 143
      ++++++ ++++++ ++++++ ++++++ ++++++
QY 200 FFLQNGYFMSNSQPMHNDSTACQFPAQIVYKTAEEQNPLQVTSRSPFSWMLSPV 259
      ++++++ ++++++ ++++++ ++++++ ++++++
DB 144 EHQUSCTWFHSHSGMSNAEAKRYQQLKNAHLVINSREKQNFVQKYL--GSATIMKGLSD- 200
      ++++++ ++++++ ++++++ ++++++ ++++++
QY 260 NQGVITWQWVQSPSPSPQKYNSSGPN--SGNEDCAFSSGSGWR 312
      ++++++ ++++++ ++++++ ++++++ ++++++
DB 201 -PTAKMKVVFVAVATSPQN-WKEDQPTWQWQVQSGHVSQDRCAPHPVQPMN 292
      ++++++ ++++++ ++++++ ++++++ ++++++
QY 413 WICK 316
      ++++++
DB 259 WICK 262

RESULT 3
US-09-978-295A-231
: Sequence 231, Application US/09978295A
: Patent No. 9820020156006A1
: GENERAL INFORMATION:
: APPLICANT: Asitkenazi, Avi
: APPLICANT: Baker Kevin P.
: APPLICANT: Bolstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroli, Ellen
: APPLICANT: Fon, Sherman
: APPLICANT: Fon, Sherman

APPLICANT: Gao, Wei-Guang
APPLICANT: Gerber, Hanspeter
APPLICANT: Herritsen, Mary E.
APPLICANT: Hoddard, Audrey
APPLICANT: Hoddowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Jarney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavitt, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Rapier, Mary A.
APPLICANT: Fan, James
APPLICANT: Facchi, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Williams, F. M. Key
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleo
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PCL1
CURRENT APPLICATION NUMBER: US/09/978,295A
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/069304
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077459
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078094
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
```


RESULT 4
US-09-978-97-231
Sequence 231, Application US/09978647
Patent No. US20020169284A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Baton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Herber, Hanspeter
APPLICANT: Gertlson, Mary F.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kiljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleotides
FILE REFERENCE: P26391027
CURRENT APPLICATION NUMBER: US/09/978,697
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/218,085
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/062,250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064,249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065,311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066,364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077,450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077,632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077,641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077,649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077,791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078,004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078,886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078,936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078,910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078,939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079,294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079,556
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079,664
PRIOR FILING DATE: 1998-03-27

PRIOR APPLICATION NUMBER: 60/079,689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079,663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079,728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079,786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079,920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079,923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080,105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080,107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080,165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080,194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080,327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080,328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080,333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080,334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081,070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081,049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081,071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081,195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081,203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081,229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081,955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081,817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081,819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081,952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081,838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082,568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082,569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082,704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082,804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082,700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082,797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082,796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083,336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083,322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083,392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083,495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083,496

```

PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/081499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/081545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/081554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085679
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
:
Query Match      17.1%   Score 289, 5%, DB 9, Length 293;
Best Local Similarity 29.4%   Pred No. 7, 1e-16;
Matches 91; Conservative 47; Mismatches 11; Indexes 61, Gaps 13.

UY      23 LVIGLI.SFMILACVAILVOVSQKPPSSLSDBDSYUAIYNNLJYKAANGELSEK... 78
DB      34 LALAIVITTVIAMVILISL-LSKASTERAALLDGHLLRTNASKOJTAALGALKREVGDC 91
UY      79 -SKLOEIVOELOTLKAANGELPEKSLOEIVOELTRIKAANGELPEKSLOEIVOELTR 137
DB      92 HSCSGSQAOQLOTTRALEJC--AQAKIME-----OSASLEIRERVTQG 133
UY      138 KAANGELPEKSLOEIVOELOTLKAANGELPEKSLOEIVOELOTLKIATFEFLT-RHFKR 197
DB      134 LAEGAGRE-----DVATELRALAEAV-RLONNS-----CEPCTPS 168

```

```

OY      198 WITFOENCGEMNSORRNHSDVIAQCHPRAUVAIETAEQNPLDLOJTSRSH
|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
LB      169 WLSFEGSCFFSEVFERTTAAAGCCHCAALASMLHYTVSGLDEGCFILFNTH
|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
OY      258 DLNQEST...WOWMDCSFLSLSPFYNNSEFPNNS-IGNETCAEPFGSG-WK
|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
LB      237 AVPHLEKVGOGYGMWGVSLSTYS---HWNGCEPPNDAMTPENCYMHLLHTGLWM
|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
OY      312 YWICKRPAC 321
|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
LB      284 GWICEKHHC 293
|.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|

RESULT 5
US 09 978 192A-231
Sequence 231, Application HS/09978192A
Patent No. US202017753A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnuyers, Luc
APPLICANT: Ealoch, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvarcoff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Gudowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavln, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paull, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tunas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypept...
FILE REFERENCE: P2630PIC9
CURRENT APPLICATION NUMBER: US/09/978, 192A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/064304
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886

```

PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078444
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078410
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079954
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079489
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079746
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079924
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081917
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081919
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081938
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082706
PRIOR FILING DATE: 1998-04-22

PRIOR APPLICATION NUMBER: 60/082747
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082746
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083346
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083406
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083544
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 17.1% Score 289.5; DB 9; Length 293;
Genetic Similarity 29.1%, Pred. No. 7, le-16;
Matches 91, Conservative 47, Mismatches 111, Indels 61, Gaps 13;

```

QY 23 LVLQLLSLMLAGVLAIVLVQVSKVPSLSLQSECEUAIYONLITQLKAAGELSEK-----78
DB 34 LALAIVLTVLMAVILITL--LSKASTERAALLDGHDLRTNSKOTAAIGALKEEVGAC 91
QY 79 -SKIAEIVLDEL PULKAAGVLELPKSKIGELYLHITLKAAGVLELPKSKIGELYLHITL 137
DB 92 HSNQSTIQAGLQITPAELTSE--AAKALME-----GNSALREIPEVYTG 133
QY 138 KAAGVLELPKSKIGELYLHITLKAAGVLELPKSKIGELYLHITLQKAEKLCRCPD 197
DB 134 LAAGGKRE-----DVRELEFALBAV-RLONN-----CECPPTS 168
QY 198 WFFPUNVYPMNSQJPMNHQSVTAQVEVKAUUVIKIAPANEIOLQSPSNFFSWGLS 257
DB 169 WISFESQYFFVYPKITMAALQIHATASAHVIVVQILQELTETKAT--KSGVWGLP 226
QY 258 WNMOST---WQVWQSPLESPSPQYWNSSPENN--NNEPCAFESSSG-WNNPCTVFN 311
DB 227 AVPHLEKVDQYQWVDQVSLSPS---HMQCFEINDAWPENCVMILHTGLMNDAPQDSFKD 283
QY 312 YWICKKPAAC 321
DB 284 CWICEKKNH 293

```

```

RESULT 5
US-09-864-761-40945
; Sequence 40945, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hank, David P.
; APPLICANT: Hancel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENE-ME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Accm13-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/754,595
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 68 24263.5
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/246,454
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001 01 30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 60/508,408

```

```

; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 40945
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC038812.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADIP T-LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEPA1, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN PONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: P26717, EVALU6 6 00e-0
; OTHER INFORMATION: EST_HUMAN HIT: 60668 1, EVALU6 2 00e-0
US-09-864-761-40945

```

```

Query Match 16.7%; Score 284; DB 10; Length
Best Local Similarity 108.3%; Hitd. No. 2; De 16;
Matches 50; Conservative 0; Mismatches 0; Indels 0;

```

```

QY 189 FLCHHCPEKIMTFEPCNCFPMNSQJPMNHQSVTAQVEVKAUUVIKIAPANEIOLQSPSNFFSWGLS
DB 1 FLCHHCPEKIMTFEPCNCFPMNSQJPMNHQSVTAQVEVKAUUVIKIAPANEIOLQSPSNFFSWGLS

```

```

RESULT 7
US-09-764-870-279
; Sequence 279, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: NO-1-6, A-13s, F-13-13s, and Antibody
; FILE REFERENCE: PTL24
; CURRENT APPLICATION NUMBER: US/09/764,870
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION DATA REMOVED - CONSULT PALM OF FILE 0793
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 279
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-279

```

```

Query Match 16.2%; Score 274.5; DB 10; Length
Best Local Similarity 27.2%; Hitd. No. 1; De 14;
Matches 81; Conservative 51; Mismatches 123; Indels 12;

```

```

QY 36 ULVALIVVQVSKVPSLSLQSECEUAIYONLITQLKAAGVLELPKSKIGELYLHITL
DB 58 LVAVALIVVQVSKVPSLSLQSECEUAIYONLITQLKAAGVLELPKSKIGELYLHITL
QY 96 GELP-EKSKIGELYLHITLKAAGVLELPKSKIGELYLHITLKAAGVLELPKSKIGELYLHITL
DB 112 VQIVQVNSQJPMNHQSVTAQVEVKAUUVIKIAPANEIOLQSPSNFFSWGLS
QY 150 LQFVLEIETLKAAGVLELPKSKIGELYLHITLKAAGVLELPKSKIGELYLHITL
DB 156 LKASALNLIKIALGSLNLSMNSHLEPQNTITQVNS-----GQVY
QY 206 YPMNSQJPMNHQSVTAQVEVKAUUVIKIAPANEIOLQSPSNFFSWGLS
DB 212 YPFLIPKTIWYSAQPPVSPNSHITSVTSESPQEF--LYKTAISGLTWIG
QY 266 QWVWQSPLESPSPQYWNSSPENN--NNEPCAFESSSG-WNNPCTVFN
DB 276 SWVWQSPLESPSPQYWNSSPENN--NNEPCAFESSSG-WNNPCTVFN

```

```

RESULT 8
US-09-905-291A-377
: Sequence 377, Application US/09905291A
: Patent No. US20020160474A1
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Rotstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Paton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gertlisen, Mary E.
: APPLICANT: Goddard, A.
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Gunney, Austin L.
: APPLICANT: Hillan, Kenneth, J.
: APPLICANT: Kiljavin, Ivar J.
: APPLICANT: Mather, Jennie P.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Thomas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William, I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: 10466-14
: CURRENT APPLICATION NUMBER: US/09/905,291A
: PRIORITY FILING DATE: 2001-07-12
: PRIOR APPLICATION NUMBER: PCT/US99/04414
: PRIORITY FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: US 60/143,048
: PRIORITY FILING DATE: 1999-07-07
: PRIOR APPLICATION NUMBER: US 60/145,698
: PRIORITY FILING DATE: 1999-07-26
: PRIOR APPLICATION NUMBER: US 60/146,222
: PRIORITY FILING DATE: 1999-07-28
: PRIOR APPLICATION NUMBER: PCT/US99/20594
: PRIORITY FILING DATE: 1999-09-08
: PRIOR APPLICATION NUMBER: PCT/US99/20944
: PRIORITY FILING DATE: 1999-09-13
: PRIOR APPLICATION NUMBER: PCT/US99/21090
: PRIORITY FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/21547
: PRIORITY FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/22089
: PRIORITY FILING DATE: 1999-10-05
: PRIOR APPLICATION NUMBER: PCT/US99/28014
: PRIORITY FILING DATE: 1999-11-29
: PRIOR APPLICATION NUMBER: PCT/US99/28313
: PRIORITY FILING DATE: 1999-11-30
: PRIOR APPLICATION NUMBER: PCT/US99/28564
: PRIORITY FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US99/28565
: PRIORITY FILING DATE: 1999-12-02
: PRIOR APPLICATION NUMBER: PCT/US99/40066
: PRIORITY FILING DATE: 1999-12-16
: PRIOR APPLICATION NUMBER: PCT/US99/40911
: PRIORITY FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US99/40949
: PRIORITY FILING DATE: 1999-12-20
: PRIOR APPLICATION NUMBER: PCT/US99/40974
: PRIORITY FILING DATE: 2000-01-05
: NUMBER OF SEQ ID NOS: 423
: SEQ ID NO 377

```

```

: LENGTH: 219
: TYPE: PRT
: ORGANISM: Homo Sapien
US-09-905-291A-377

Query Match: 16.1% Score 274.5; DB 9; Length 219;
Best Local Similarity 38.5% Pred. No. 9, 1e-15;
Matches 60; Conservative 25; Mismatches 56; Indels 15; Gaps 6;

QY 172 KGGGVVPELTHF-----IAPELVPHVYKDWVPHQGNVYFMSNSQPMWHSVYVQGVFV 226
: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 55 KKEOLFENFETLSYNGSGSVKNC--CPINWEYFGSSCYFSIDTISWALSLKNCAMG 112
: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 227 AGIVVYKFAFEGNITQIQISQSNQPMWMTSPVNDQGTWQWVGSPISSPGRVWNSGTP 286
: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 113 AHIVVINSQFFQFPHSYKPKKMPF-FVGLSQGVVGSQMWVQTPVFKSLIS-FWVQVGP 170
: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 287 NNSGN-ELCAEFSNS-----GWDNKKQVDVNWVCK 316
: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 171 NNIALEDCATMKDSSNPQGNMNDVTCPLNYPRTCE 206
: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 9
US-09-909-320-377
: Sequence 377, Application US/09909320
: Patent No. US20020132240A1
: GENERAL INFORMATION:
: APPLICANT: Genentech, Inc.
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Rotstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Paton, Dan L.
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Gerber, Hanspeter
: APPLICANT: Gertlisen, Mary E.
: APPLICANT: Goddard, A.
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, Christopher J.
: APPLICANT: Gunney, Austin L.
: APPLICANT: Hillan, Kenneth, J.
: APPLICANT: Kiljavin, Ivar J.
: APPLICANT: Mather, Jennie P.
: APPLICANT: Pan, James
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Thomas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William, I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: FILE REFERENCE: 10466-14
: CURRENT APPLICATION NUMBER: US 60/146,222
: PRIORITY FILING DATE: 2002-01-04
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIORITY FILING DATE: 2000-02-22
: PRIOR APPLICATION NUMBER: US 60/143,048
: PRIORITY FILING DATE: 1999-07-07
: PRIOR APPLICATION NUMBER: US 60/145,698
: PRIORITY FILING DATE: 1999-07-26
: PRIOR APPLICATION NUMBER: US 60/146,222
: PRIORITY FILING DATE: 1999-07-28
: PRIOR APPLICATION NUMBER: PCT/US99/20594
: PRIORITY FILING DATE: 1999-09-08
: PRIOR APPLICATION NUMBER: PCT/US99/20944
: PRIORITY FILING DATE: 1999-09-13
: PRIOR APPLICATION NUMBER: PCT/US99/20944
: PRIORITY FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/21090
: PRIORITY FILING DATE: 1999-09-15
: PRIOR APPLICATION NUMBER: PCT/US99/21547
: PRIORITY FILING DATE: 1999-09-15

```


APPLICANT: Pan,James
APPLICANT: Smith,Victoria
APPLICANT: Matanabe,Colin K.
APPLICANT: Wood,William L.
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430101
CURRENT APPLICATION NUMBER: 08/10/052,596
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062450
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/064121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063440
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066446
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078839
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333

PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05


```
? PRIOR APPLICATION NUMBER: 60/088202
? PRIOR FILING DATE: 1998-06-05
? PRIOR APPLICATION NUMBER: 60/088212
? PRIOR FILING DATE: 1998-06-05
? PRIOR APPLICATION NUMBER: 60/088217
? PRIOR FILING DATE: 1998-06-05
? PRIOR APPLICATION NUMBER: 60/088326
? PRIOR FILING DATE: 1998-06-04
? PRIOR APPLICATION NUMBER: 60/088655
? PRIOR FILING DATE: 1998-06-09
? PRIOR APPLICATION NUMBER: 60/088722
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088738
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088743
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088811
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088824
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088825
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088826
? PRIOR FILING DATE: 1998-06-10
? PRIOR APPLICATION NUMBER: 60/088861
? PRIOR FILING DATE: 1998-06-11
? PRIOR APPLICATION NUMBER: 60/088863
? PRIOR FILING DATE: 1998-06-11
? PRIOR APPLICATION NUMBER: 60/088876
? PRIOR FILING DATE: 1998-06-11
? PRIOR APPLICATION NUMBER: 60/089090
? PRIOR FILING DATE: 1998-06-12
? PRIOR APPLICATION NUMBER: 60/089105
? PRIOR FILING DATE: 1998-06-12
? PRIOR APPLICATION NUMBER: 60/089512
? PRIOR FILING DATE: 1998-06-16
? PRIOR APPLICATION NUMBER: 60/089514
? PRIOR FILING DATE: 1998-06-16
? PRIOR APPLICATION NUMBER: 60/089538
? PRIOR FILING DATE: 1998-06-17
? PRIOR APPLICATION NUMBER: 60/089598
? PRIOR FILING DATE: 1998-06-17
? PRIOR APPLICATION NUMBER: 60/089653
? PRIOR FILING DATE: 1998-06-17
? PRIOR APPLICATION NUMBER: 60/089908
? PRIOR APPLICATION NUMBER: 60/089908

Query Match 16.1% Score 273.5; DB 12; Length 219;
Best Local Similarity 38.5%; Pred. No. 9; Indels 15; Gaps 6;
Matches 60; Conservative 25; Mismatches 56;
```

```
? 172 KQOQYVQELTTLK...TAFELTPEHCPKILW:FGUNYHMSNSGNNMH:SVTA:GEV: 27%
? 55 KKFQLEPNEFTLSCYVNSGSVKNC-CPLNWEYFQSSCYFFSDTISMLSLKNS:AMG 112
? 227 AOLVYIKIAEKNF:QLOTSRNSRNSWMLGSLNQBGTWQVWDSPI:SPSPQRYWNSGEP 286
? 113 AHLVYINSOEBOEFLSYKKPKMREF-FLGLSDYVEGQWQWVDGDTPLTKSIS-FWDVGEIP 170
? 287 NNSGN-EDCAEFGS-----GMDNRCEDVDNYWICK 316
? 171 NNIALIHOCATIRROSSNRRKMMNVDYTCLELNFRIICE 206

RESULT 12
US-09-862-802-5
? Sequence 5, Application US/09862802
? Patent No. US20020165346A1
? GENERAL INFORMATION:
? APPLICANT: Schering-Plough Corporation
? TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
? FILE REFERENCE: SP0695B
? CURRENT APPLICATION NUMBER: US/09/862.802
? CURRENT FILING DATE: 2002-01-10
```

```
? PRIOR APPLICATION NUMBER: US 09/111,470
? PRIOR FILING DATE: 1998-07-04
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: Patent in version 3.1
? SEQ ID NO 5
? LENGTH: 291
? TYPE: PRT
? ORGANISM: Unknown
? FEATURE:
? OTHER INFORMATION: mammalian protein
US-09-862-802-5
```

```
Query Match 16.0% Score 271.5; DB 9; Length
Best Local Similarity 27.0%; Pred. No. 1; Indels 14;
Matches 86; Conservative 45; Mismatches 112;
```

```
? 6 EPRVQGLGIGLGGALVQLSLFMILAVIVALLVQSVSPSLSGEG
? 29 QPILORL-----CGPRLLILSL-----GLSLILVAVVIGSUNSLG
? 66 TLKAAVGLSEKSKTIGTYQPI:TLKAAV:PIPKSKLQETLYETL:K
? 78 SNTFAST-----FAVKEISTGNNVGRKMSLE-----
? 126 KLOEYQELTLKAAVGLPEKSKLQETLYQEL:TLKAAVGLPIQSKQGG
? 108 QLEKQKDLSE-----DHSST-----LHVKQFVSDLRSLISQMA
? 186 AFEELCPHCPKILW:FGUNYHMSNSGNNMH:SVTA:GEV:AVQI:VLR:IA
? 148 GSERIC-CPVWNVHENSCTWPSRSKAMALANVYCLDAHLVAVTSW
? 246 SPSNPSWMLGSLNQBGTWQVWDSPI:SPSPQRYWNSGEPNN-----S
? 206 GFVN-TWMLHD-QNGPWKQWVDGDIYETSPKN-WPEQDDWYGHIL
? 300 SG-WINDNCDVDNYWICK 316
? 261 DQRMNDVQRPYRWCE 276
```

```
RESULT 13
US-09-862-802-6
? Sequence 6, Application US/09862802
? Patent No. US20020165346A1
? GENERAL INFORMATION:
? APPLICANT: Schering-Plough Corporation
? TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; REI
? FILE REFERENCE: SP0695B
? CURRENT APPLICATION NUMBER: US/09/862.802
? CURRENT FILING DATE: 2002-01-10
? PRIOR APPLICATION NUMBER: US 09/111,470
? PRIOR FILING DATE: 1998-07-04
? NUMBER OF SEQ ID NOS: 11
? SOFTWARE: Patent in version 3.1
? SEQ ID NO 6
? LENGTH: 287
? TYPE: PRT
? ORGANISM: Unknown
? FEATURE:
? OTHER INFORMATION: mammalian protein
US-09-862-802-6
```

```
Query Match 15.8% Score 267.5; DB 9; Length
Best Local Similarity 28.8%; Pred. No. 3; Indels 14;
Matches 92; Conservative 49; Mismatches 103; Indels 14;
```

```
? 6 EPRVQGLGIGLGGALVQLSLFMILAVIVALLVQSVSPSLSGEG
? 28 QPILORL-----GLSLILVAVVIGSUNSLG
? 66 TLKAAVGLSEKSKTIGTYQPI:TLKAAV:PIPKSKLQETLYETL:K
```

```

Db 70 KSLKEAFSNFS-SSTLEFV-QALSTHGGSVD-----KITSIGA----- 106
QY 126 KLOEYQHTLRKAAGSELPKSKLQELIYQELIRKAAGSELPKSKQVQYQHTLHKL 185
Db 107 KLEQVQDQ--LKADHDAL-----LPHLKHFPVDLFEVACQWELH-----SN 145
QY 186 APERLCHCHCKEDWFFPGNCFYFMSNSQPMHNSVTACQVPAQVLVIRKAEQNFILQGT 245
Db 147 QSKRTG--CPVNWVERHGTGYWFSHSHSKAMAPAKYCOLFNALIVITNSMTYQKTVQHT 294
QY 246 SESNPFESMGLSDLNQETWQWVQSPFISYFQPYNSCEFN-----SCNEDCAEFSG 299
Db 205 NPEN--TWIGITD--SDGSWKWVQGTIDYRHNRYN-WAVTQPDWGHSHPLGSGEDVEVAP 259
QY 300 SG-WNDNRQCDYDWYICK 317
Db 260 EPPNRHPPQELQVWVWVFK 298

```

```

RESULT 14
US-10-090-466-4
: Sequence 4: Application US/10090466
: Patent No. US20020137914A1
: GENERAL INFORMATION:
: APPLICANT: Turner, C. Alexander Jr.
: APPLICANT: Mathur, Brian
: APPLICANT: Gullinan, Emily B.
: TITLE OF INVENTION: No. US20020137914A1-1 Human Protein and Polynucleotide
: FILE REFERENCE: LEX-0315-USA
: CURRENT APPLICATION NUMBER: US/10/090,466
: PRIOR FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 60/274,961
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 4
: LENGTH: 182
: TYPE: PRT
: ORGANISM: homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)...(182)
: OTHER INFORMATION: Xaa - Any Amino Acid
US-10-090-466-4

```

```

Query Match 15.5%; Score 263.5; DB 12; Length 182;
Best Local Similarity 35.9%; Pred. No. 4, 4e-14;
Matches 65; Conservative 27; Mismatches 66; Indels 23; Gaps 8;

```

```

QY 144 LPEKSLQETLYQELTPRKAAVTPIPIYQSKQYLYQHT---DLKTAPEPLCPHCKWMTF 200
Db 11 VPHNPMWSKTVKRLSKLR-----EYQYHXSLLTCVMECKDLEIMSC--CPTPWTS 58
QY 201 FQGNVYFMSNSQPMHNSVTACQVPAQVLVIRKAEQNFILQGTSESNPFSMGLSDLN 260
Db 59 FQSSCYFTISTGMQSWTKSKQNFVYMGADLVVITNREDOFT-LQNKRNSSYFLGSDPG 117
QY 261 QGTWQWVQDQSPFISYFQPYNSCEFN-----GWNDRQCDYDWYIC 315
Db 118 GRHMQWVQDQPYNFNV-TFWHSGEPNNI-LDEGCATINFRSSEMGWMDIHCHVQKSLG 175
QY 316 K 316
Db 176 K 176

```

```

RESULT 15
US-10-090-466-2
: Sequence 2: Application US/10090466
: Patent No. US20020137914A1
: GENERAL INFORMATION:
: APPLICANT: Turner, C. Alexander Jr.
: APPLICANT: Mathur, Brian

```

```

: APPLICANT: Gullinan, Emily B.
: TITLE OF INVENTION: No. US20020137914A1-1 Human Protein and Polynucleotide
: FILE REFERENCE: LEX-0315-USA
: CURRENT APPLICATION NUMBER: US/10/090,466
: PRIOR FILING DATE: 2002-03-01
: PRIOR APPLICATION NUMBER: US 60/274,961
: PRIOR FILING DATE: 2001-03-12
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 2
: LENGTH: 213
: TYPE: PRT
: ORGANISM: homo sapiens
US-10-090-466-2

```

```

Query Match 15.5%; Score 263.5; DB 12; Length 213;
Best Local Similarity 45.9%; Pred. No. 5, 4e-14;
Matches 65; Conservative 27; Mismatches 66; Indels 23; Gaps 8;

```

```

QY 144 LPEKSLQETLYQELTPRKAAVTPIPIYQSKQYLYQHT---DLKTAPEPLCPHCKWMTF 200
Db 42 VPHNPMWSKTVKRLSKLR-----EYQYHXSLLTCVMECKDLEIMSC--CPTPWTS 89
QY 201 FQGNVYFMSNSQPMHNSVTACQVPAQVLVIRKAEQNFILQGTSESNPFSMGLSDLN 260
Db 90 FQSSCYFTISTGMQSWTKSKQNFVYMGADLVVITNREDOFT-LQNKRNSSYFLGSDPG 148
QY 261 QGTWQWVQDQSPFISYFQPYNSCEFN-----GWNDRQCDYDWYIC 315
Db 140 GRHMQWVQDQPYNFNV-TFWHSGEPNNI-LDEGCATINFRSSEMGWMDIHCHVQKSLG 206
QY 316 K 316
Db 207 K 207

```

```

Search completed: December 7, 2002, 11:08:16
Job time : 24 secs

```



```

CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 09/08/365,103B
: FILING DATE: 28-DEC-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Nebel, Heidi S.
: REGISTRATION NUMBER: 37,719
: REFERENCE/DOCKET NUMBER: 0111 N5-24
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (515) 288-4667
: TELEFAX: (515) 288-1348
: INFORMATION FOR SEQ. ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 300 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-365-103B-6

```

```

Query Match: 19.94; Score 338; DP 1; Length 300;
Best Local Similarity 30.49; Prod No. 2,10-23;
Matches 93; Conservative 59; Mismatches 104; Indels 50; Gaps 14.

```

```

QY 34 AGVLAIVLVQSKVPSLSLQSEFQDAIYQNTLYL... KAAVGELSEKSKIOYQEL 88
DB 11 AGLLALLIMHMETEKMLKQ---LGDIAIONVSHVTKDLQKPSNQLAKSQVYVMSQNL 67
QY 89 TOLKAAGELPEKSKLOIYQITRLKAAGELPEKSKLOIYQELTLK-----AAVG 142
DB 68 QELAEQKQK--KQDPSLSNLTGLQEDLPN--AQSNKLSQNLNLDLDDLVNIKSLG 123
QY 143 ELPEK---SKLOEIYQELTRKAAGELPDQSKQDIYQELDLTAIFERLRHFKM 198
DB 124 LNKRIASISLEKIQEYVAKLMEI...-----LISKGA-----CNLEPKMW 163
QY 169 TPEQCNVYFMSNSQPMNHDSVTAQVEYPAQIVYIKTAERPNFQICTSPNPFNMTSF 258
DB 164 LHFQOKIYVEFGKSGKQWIOARFACSDLOGPLVSHSUKEDQDFLMQHINKKD--SWGLDD 221
QY 259 INQGTQWQVNSPLSPSPYWNNSGPNNSG--NEDVAFPSGSG--WINDPG--IVIRYWG 315
DB 222 LMRERFVWSRSPVYS---NMNPGPNNGRQETL--VMNPGQWNTAFQPSYITAAWG 278
QY 416 KKPAC 321
DB 279 HOLATC 284

```

```

RESULT 5
: US-08-365-103B-4
: Sequence 4, Application US/08365103B
: Patent No. 5766943
: GENERAL INFORMATION:
: APPLICANT: Lynch, Richard G.
: APPLICANT: Nunez, Raphael D.
: APPLICANT: Yodanis, Jungi
: TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Zarley, McKee, Thomee, Voorhees & Sease
: STREET: 801 Grand Ave. Suite 3200
: CITY: Des Moines
: STATE: Iowa
: COUNTRY: United States
: ZIP: 50309
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 09/08/365,103B

```

```

FILING DATE: 28-DEC-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Nebel, Heidi S.
: REGISTRATION NUMBER: 37,719
: REFERENCE/DOCKET NUMBER: 0111 N5-24
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (515) 288-4667
: TELEFAX: (515) 288-1348
: INFORMATION FOR SEQ. ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 287 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-365-103B-4

```

```

Query Match: 19.74; Score 344; DP 1; Length 287;
Best Local Similarity 31.96; Prod No. 4,60-23;
Matches 93; Conservative 47; Mismatches 101; Indels 50; Gaps 14.

```

```

QY 44 VSKVPSLSQSEFQDAIYQNTLYL... KAAVGELSEKSKLOIYQELTLK 88
DB 26 VSHVTKDLQKPSNQLAKSQVYVMSQNL 67
QY 104 LQFIYQITRLKAAGELPEKSKLOIYQELTLKKAAGELPEK-----SGL 142
DB 75 LTGLQEDLPN--AQSNKLSQNLNLDLDDLVNIKSLG 123
QY 160 LKAAGELPEKSKLOIYQITRLKAAGELPDQSKQDIYQELDLTAIFERLRHFKM 198
DB 131 LMEI...-----LISKGA-----CNLEPKMW 163
QY 220 TACQVPAQIVYIKTAERPNFQICTSPNPFNMTSF 258
DB 170 FAFQIKIYVEFGKSGKQWIOARFACSDLOGPLVSHSUKEDQDFLMQHINKKD--SWGLDD 221
QY 280 YWNGEPPNSG NEDVAFPSGSG--WINDPG--IVIRYWG 315
DB 227 NMNPGPNNGRQETL--VMNPGQWNTAFQPSYITAAWG 278

```

```

RESULT 6
: US-09-535-521-2
: Sequence 2, Application US/09535521
: Patent No. 6410714
: GENERAL INFORMATION:
: APPLICANT: Weber, Eric R.
: APPLICANT: McCall, Catherine A.
: TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR
: FILE REFERENCE: AL-5
: CURRENT APPLICATION NUMBER: US/09-535,521
: EARLIER APPLICATION NUMBER: 60/125,913
: EARLIER FILING DATE: 1999-04-24
: NUMBER OF SEQ. ID NOS: 26
: SOFTWARE: Patent In Ver. 2.1
: SEQ. ID NO. 2
: LENGTH: 292
: TYPE: PRT
: ORGANISM: Canis familiaris
: US-09-535-521-2

```

```

Query Match: 19.48; Score 329.5; DP 4; Length 292;
Best Local Similarity 29.55; Prod No. 1,20-23;
Matches 97; Conservative 45; Mismatches 102; Indels 50; Gaps 14.

```

```

QY 11 QLGILGTAAGAVIGLSPMLAGVVAHVYVSKVPSLSQSEFQDAIYQNTLYL... KAAVGELSEKSKLOIYQELTLK 88
DB 27 QLAGLG---LVIVMAGLITLILFW-----HRI 67
QY 70 AAVPTSEKSKLOIYQITRLKAAGELPDQSKQDIYQELDLTAIFERLRHFKM 198

```

```

Db 63 AAAGNSRVSCKDLERHNG-----DQMAKQSAQAQVSOD-----MKE 98
QY 130 IY-ELLILKAAGVELPEKSKLOEYIYELLIFLKAAGVELFGVS--KQCCYQYELMLKLA 187
Db 99 IOAEOKHMKADSELS-----ONLDALRSDLNKLSQSLNTERSTALHLEKLEFV 149
QY 188 EPI-----CRICIKDWITFPQCNCTPMNSQBNHDSYTAQOEALVYIKTAE 236
Db 150 FKIMELHVSNGSEQNTCPKRNLPQKTYFGSEPEKKMIQAFAPSKIQPLASHSGE 209
QY 237 EGNELQIOTSRNSNFEFMMGLSDLNQCTWQWVDSPLSPSFQYVNSGEPNNSQ-NEDEA 295
Db 210 EODEFLAYANKKK--TWIGLRDLPDPEFTIMDPNPLNVS--NMPEPEPNNSQGPPIGV 264
QY 296 EFSNSG-WNNRPG--EVDNWIQKKAAC 321
Db 265 MMQTSQWMDNAPFGSSLDG-WVWDPLATG 292

```

```

RESULT 7
US-09-535-521-5
: Sequence 5, Application US/09535521
: Patent No. 6410714
: GENERAL INFORMATION:
: APPLICANT: Weber, Eric R.
: APPLICANT: McCall, Catherine A.
: TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
: TITLE OF INVENTION: PROTEINS, NOVELLY ACID MOLECULES AND USES THEREOF
: FILE REFERENCE: AL-5
: CURRENT APPLICATION NUMBER: US/09/535,521
: CURRENT FILING DATE: 2000-03-24
: EARLIER APPLICATION NUMBER: 60/125,913
: EARLIER FILING DATE: 1999-03-24
: NUMBER OF SEQ. ID NOS.: 26
: SOFTWARE: patentin Ver. 2.1
: SEQ. ID NO. 5
: LENGTH: 292
: TYPE: PRI
: ORGANISM: Canis familiaris
US-09-535-521-5

```

```

Query Match: 19.4%, Score 329.5, DB 1, Length 292;
Best Local Similarity 29.5%, Prod. No. 1.2e-22;
Matches 92; Conservative 49; Mismatches 102; Indels 81; Gaps 15;

QY 11 QSLILCLGAGALVLIQSLFPMILAGVVALVQVSVPSLSQESQEPATVQNTQK 69
Db 27 QSLILCLGAGALVLIQSLFPMILAGVVALVQVSVPSLSQESQEPATVQNTQK 69
QY 70 AAVGELSEKSKLOEYIYELLIFLKAAGVELPEKSKLOEYIYELLIFLKAAGVELPEKSKLOE 129
Db 64 AAVGELSEKSKLOEYIYELLIFLKAAGVELPEKSKLOEYIYELLIFLKAAGVELPEKSKLOE 129
QY 130 IY-ELLILKAAGVELPEKSKLOEYIYELLIFLKAAGVELFGVS--KQCCYQYELMLKLA 187
Db 99 IOAEOKHMKADSELS-----ONLDALRSDLNKLSQSLNTERSTALHLEKLEFV 149
QY 188 EPI-----CRICIKDWITFPQCNCTPMNSQBNHDSYTAQOEALVYIKTAE 236
Db 150 FKIMELHVSNGSEQNTCPKRNLPQKTYFGSEPEKKMIQAFAPSKIQPLASHSGE 209
QY 237 EGNELQIOTSRNSNFEFMMGLSDLNQCTWQWVDSPLSPSFQYVNSGEPNNSQ-NEDEA 295
Db 210 EODEFLAYANKKK--TWIGLRDLPDPEFTIMDPNPLNVS--NMPEPEPNNSQGPPIGV 264
QY 296 EFSNSG-WNNRPG--EVDNWIQKKAAC 321
Db 265 MMQTSQWMDNAPFGSSLDG-WVWDPLATG 292

```

RESULT 8
US-08-465-1048-8

```

: Sequence 8, Application US/08465103B
: Patent No. 5766943
: GENERAL INFORMATION:
: APPLICANT: Lynch, Richard G.
: APPLICANT: Nam-2, Raphael D.
: APPLICANT: Yodot, Jungi
: TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
: TITLE OF INVENTION: and Methods of Use for Same
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Farley, McKee, Thompson, Voorhees & Sease
: STREET: 801 Grand Ave., Suite 3200
: CITY: Des Moines
: STATE: Iowa
: COUNTRY: United States
: ZIP: 50309
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 08/047365,103B
: FILING DATE: 28-DEC-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Nobel, Heidi S.
: REGISTRATION NUMBER: 37,719
: PREPENDING/CHECKER NUMBER: 0115 N5 24
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (515) 288-3667
: TELEFAX: (515) 288-1338
: INFORMATION FOR SEQ. ID NO.: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 321 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-465-1048-8

```

```

Query Match: 19.0%, Score 322.5, DB 1, Length 321;
Best Local Similarity 28.6%, Prod. No. 6.3e-22;
Matches 92; Conservative 61; Mismatches 121; Indels 55; Gaps 13;

QY 1 MSLSKEFEVQGLLQVLIQSLFPMILAGVVALVQVSVPSLSQESQEPATVQNTQK 69
Db 1 MSLSKEFEVQGLLQVLIQSLFPMILAGVVALVQVSVPSLSQESQEPATVQNTQK 69
QY 58 AAVGELSEKSKLOEYIYELLIFLKAAGVELPEKSKLOEYIYELLIFLKAAGVELPEKSKLOE 129
Db 61 AAVGELSEKSKLOEYIYELLIFLKAAGVELPEKSKLOEYIYELLIFLKAAGVELPEKSKLOE 129
QY 117 AAVGELSEKSKLOEYIYELLIFLKAAGVELPEKSKLOEYIYELLIFLKAAGVELPEKSKLOE 172
Db 105 AAVGELSEKSKLOEYIYELLIFLKAAGVELPEKSKLOEYIYELLIFLKAAGVELPEKSKLOE 172
QY 173 QNQLQIOTSRNSNFEFMMGLSDLNQCTWQWVDSPLSPSFQYVNSGEPNNSQ-NEDEA 295
Db 150 FKIMELHVSNGSEQNTCPKRNLPQKTYFGSEPEKKMIQAFAPSKIQPLASHSGE 209
QY 237 EGNELQIOTSRNSNFEFMMGLSDLNQCTWQWVDSPLSPSFQYVNSGEPNNSQ-NEDEA 295
Db 210 EODEFLAYANKKK--TWIGLRDLPDPEFTIMDPNPLNVS--NMPEPEPNNSQGPPIGV 264
QY 296 EFSNSG-WNNRPG--EVDNWIQKKAAC 321
Db 265 MMQTSQWMDNAPFGSSLDG-WVWDPLATG 292

```

RESULT 9
US-09-111-470-4
: Sequence 4, Application US/09111470
: Patent No. 6277959

```

GENERAL INFORMATION:
APPLICANT: Valladeau, Jenny
APPLICANT: Ravel, Odile
APPLICANT: Bates, Elizabeth E.M.
APPLICANT: Ford, John
APPLICANT: Saeland, Sem
APPLICANT: Lebecque, Serge J.E.
TITLE OF INVENTION: Mammalian Membrane Protein Genes;
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,470
FILING DATE: 08-JUL-1994
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/053,080
FILING DATE: 09-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SP0645
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)452-9146
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 316 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-111-470-4

Query Match 18.8%; Score 319.5; DB 4; Length 316;
Best Local Similarity 28.8%; Pred. No. 1,26-21;
Matches 90; Conservative 53; Mismatches 103; Indels 67; Gaps 13;

QY 23 LVQLLSFMLAGVLAIVLVSVKSSLSQSESDAIYONLQLDKAAVGLSEKSKLV 82
DB 41 LLLSLGLLLIVICVGVGFSKFRDLVLRITD...FSNFIISNIVA-----ELQ 88
QY 83 ELYQELTOLKAAVGLPEKSKLQELLYELLTKAAVGLPEKSKLQELLYELLTKAAV 142
DB 89 AITSSQSSLEETIASL--KAEVEGFKDE--RQAGVSEL-----DEHTIQKALIG 133
QY 143 ELPE-----KSKLQELLYELLTKAAVGLPEKSKLQELLYELLTKAAVGLPEK 190
DB 134 HCPHCPSPVCVPHSEMLLRVQVQIVODLKLKLCQVATLNNNA-----STFGT 179
QY 191 CHHCKRDMTFPGQNTYFMSNSIPRNHHISVLAQGVPAQIVLIKAFQNTLYQLSPSNP 250
DB 180 Q--CPVNVVHEHDSQCYVSHSISMVAEKEKYQLKNAHIVVINSFPELNFQVYV--GSA 236
QY 251 FSWMGISLDLNGGTQWQWDGSLSPSFQRYNNNSFEPNN-----SGNDFAEESGNSGWN 303
DB 246 YTMMLTISL--PPTAKKWKVCTIVATCPQN--WRPQPPHWWYTHGICCTPTAAHHHTGPMN 292
QY 304 IAKKCDVDNWK 316
DB 293 DDVCORPHYWCE 305

```

```

RESULT 10
US-08-365-103B-10
Sequence 10, Application US/08-365103B
Patent No. 5766943
GENERAL INFORMATION:
APPLICANT: Lynch, Richard G.
APPLICANT: Nunez, Raphael D.
APPLICANT: Yodanis, Daniel
TITLE OF INVENTION: DNA Sequences for Soluble Forms of
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Zaffey, McKee, Thumte, Voorhees & Sease
STREET: 801 Grand Ave, Suite 1200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/365,103B
FILING DATE: 28-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: 01RT N5-24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 288-9667
TELEFAX: (515) 288-1348
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-365-103B-10

Query Match 18.7%; Score 317.5; DB 1; Length
Best Local Similarity 29.0%; Pred. No. 1,86-21;
Matches 94; Conservative 59; Mismatches 116; Indels
QY 9 VOQLGLGCIQGHCA--LVQLLSFMLAGVLAIVLVSVKSSLSQ--EDQ 10
DB 8 IELPLRRKQCPKQJIVLIGIVIAALWAGLLILLIMWMTQSLKQLEDE 10
QY 66 TOLKAAVGLPEKSKLQELLYELLTKAAVGLPEKSKLQELLYELLTKAAVGLPEK 142
DB 68 KLESIHGDQMAOKSOSDIOSELELKA-----EUGR 142
QY 125 SKIQLLYELLTKAAVGLPEKSKLQELLYELLTKAAVGLPEKSKLQELLYELLTKAA 190
DB 110 WNLNGVGLDSSEKSG--ELNEPNEASDLLEPLREVLKLR----- 142
QY 181 TOLKTAFFPLPHCPKRWTFPGQNTYFMSNSIPRNHHISVLAQGVPAQIVLIKAFQNT 250
DB 149 MFLVNSQVWVNIQELKWIWVQVQVSIQKQTEWVHAPVYAGRDMDVQV 236
QY 241 LSLQSPSEFMSMGISLDLNGGTQWQWDGSLSPSFQRYNNNSFEPNN-----SGNDF 303
DB 209 LTKHASHIS--SWIGLRNLILKGERIWDGSHVDYS--NNAPELPSKRS 292
QY 300 SG--WNNPQVD--NTWICKKPAAG 321
DB 264 SGRWNDAPCDRLKLGAWVCDRLATG 287

```

RESULT 11
US-09-111-470-10

Sequence 10, Application US/09111470
Patent No. 6277959
GENERAL INFORMATION:
APPLICANT: Valladeau, Jenny
APPLICANT: Kavel, Odile
APPLICANT: Hates, Elizabeth E.M.
APPLICANT: Ford, John
APPLICANT: Saeland, Sem
APPLICANT: Lebecque, Serge J.F.
TITLE OF INVENTION: Mammalian Membrane Protein Genes,
TITLE OF INVENTION: Related Reagents
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-POS/MS-POS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/111,470
FILING DATE: 08-JUL-1998
CLASSIFICATION:
Prior Application DATA:
APPLICATION NUMBER: US 60/053,080
FILING DATE: 09-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: SF0695
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)446-1200
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 273 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-111-470-10

Query Match: 18.08; Score 305; DB 4; Length 273;
Best Local Similarity: 29.38; Prod No 2,1e-20;
Matches: 89; Conservative: 49; Mismatches: 96; Indels: 70; Gaps: 15;

QY 20 HGAIVGLISPMITAGVIVALLVWVSKVPSISQSESDDAIYQNTOLKAVGELSEK 79
DB 22 NCPILQSL--LLVILICVGVQNSKRFQRIYLRID--FSNIT .. SN 63
QY 80 KLOFVYQELTOLKAVGELPEKSKLOETIYQELTRKAVGELPEKSKLOETIYQELTRKA 139
DB 64 TVALT-QALIS-----QGSLEE--TISLKAEEGF--KQKRAVISEML----- 104
QY 140 AVGELPEKSKLOETIYQELTRKAVGELPDQSKQOQIYQELTDLKATFELRHCKRDMT 199
DB 105 -----LPVQIVQIKKLTQVATLNNNGE-----ASTEGTQ--CIVNNV 143
QY 200 EFGNCTFFMSNSQRNMDSTAVQVEVAQVLVKTAEQNFLOLOTSRKNSRWGLSLD 259
DB 144 EHQSCVWFHSQMSMAEAKVYQIKNAHLVINSREQNFVQKYL--GSATVMGLSD-- 200
QY 260 NQPTFMQVNDSPISPFQRYWNSGEPNN-----SSENGCAFSSGS--WNINPCGVNDV 312
DB 201 -PQAKWVNDSPISPFQRYWNSGEPNN-----SSENGCAFSSGS--WNINPCGVNDV 312
QY 413 WICK 316
DB 259 WICE 262

RESULT 12
US-08-688-442-4
Sequence 4, Application US/08688442
Patent No. 5821964
GENERAL INFORMATION:
APPLICANT: Au-Yang, Janice
APPLICANT: Cocks, Benjamin G.
APPLICANT: Coll, Surya K.
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSER: Jucyle Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/688,442
FILING DATE: Filed Herewith
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0005-1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1245724
US-08-688-442-4

Query Match: 17.78; Score 303.5; DB 2; Length 292;
Best Local Similarity: 28.68; Prod No 3.1e-20;
Matches: 90; Conservative: 54; Mismatches: 98; Indels: 73; Gaps: 16;

QY 26 HGAIVGLISPMITAGVIVALLVWVSKVPSISQSESDDAIYQNTOLKAVGELSEK 68
DB 22 NCPILQSL--LLVILICVGVQNSKRFQRIYLRID--FSNIT .. SN 63
QY 80 KLOFVYQELTOLKAVGELPEKSKLOETIYQELTRKAVGELPEKSKLOETIYQELTRKA 139
DB 64 TVALT-QALIS-----QGSLEE--TISLKAEEGF--KQKRAVISEML----- 104
QY 140 AVGELPEKSKLOETIYQELTRKAVGELPDQSKQOQIYQELTDLKATFELRHCKRDMT 199
DB 105 -----LPVQIVQIKKLTQVATLNNNGE-----ASTEGTQ--CIVNNV 143
QY 200 EFGNCTFFMSNSQRNMDSTAVQVEVAQVLVKTAEQNFLOLOTSRKNSRWGLSLD 259
DB 144 EHQSCVWFHSQMSMAEAKVYQIKNAHLVINSREQNFVQKYL--GSATVMGLSD-- 200
QY 260 NQPTFMQVNDSPISPFQRYWNSGEPNN-----SSENGCAFSSGS--WNINPCGVNDV 312
DB 201 -PQAKWVNDSPISPFQRYWNSGEPNN-----SSENGCAFSSGS--WNINPCGVNDV 312
QY 413 WICK 316
DB 259 WICE 262

Db 267 WNDVCOPIHWCE 281

RESULT 13

US-09-113-788-4
Sequence 4: Application US/09113788

Patent No. 5965104

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice

APPLICANT: Cooks, Benjamin G

APPLICANT: Goll, Surya K.

APPLICANT: Hillman, Jennifer L.

TITLE OF INVENTION: NOVEL HUMAN C-TYPE LECTIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Inocyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSO Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/113,788

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/668,342

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0095-1 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-845-4166

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 292 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 123724

US-09-113-788-4

Query Match 17 99: Score 303.5: DB 2: Length 292.

Best Local Similarity 28.68: Pred No. 3.1e-20:

Matches 90: Conservative 54: Mismatches 98: Indels 73: Gaps 16:

20 HCAVLDTL-----STMLA---GVAVILVQVSKVSSLSQSEDAIYQNLVL 68

22 NGLPIQSLQRLRSGPCHLISLGIILIVIGVVGQNSKFG-----PDLVTL 72

69 KAAVGLSEKSKLOEYQELTQKAAVGLPEKSKLOEYQELTQKAAVGLPEKSKLTG 128

74 KIDPSNTNST-VAEI-QALTS-----QSSLEF--TASIKAVESEF--KQERQ 116

129 ELYQELTRKAAVGLPEKSKLOEYQELTQKAAVGLPEKSKLOEYQELTQKAAVGLPEKSKLTG 188

117 AVHSEML-----LRVQQLVDLKLKLTQVATLNNNEE-----ASTE 153

189 KLRCHQKDKMTFFQGNFYFMSNSQRNMHDSVTAQCEVRKQVLYVYIKTAQONFLQLOTSS 248

154 GTC--CVNNVNEHDSQYWFSSHGSMALAEKYQCLNHLVINSREKQNFVQKYL--C 209

249 NRESWMLSDLNQGTQWQVDAISPLSPFQRYVNSGEFNN-----STNFTFAEFSGSF- 301

Db 210 SAYTWMLESD--PRGANKWVETIIVAIIDQGN-WKPSGJHQQWQGHQI QKQDEP
QY 302 WNDNKCVDNMYWICK 316
Db 267 WNDVCOPIHWCE 281

RESULT 14

US-09-535-521-20
Sequence 20: Application US/0955521

Patent No. 6410714

GENERAL INFORMATION:

APPLICANT: Weber, Eric R.

APPLICANT: McCall, Catherine A.

TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR

FILE REFERENCE: AL-5

CURRENT AFFILIATION NUMBER: 5270-0545, 521

CURRENT FILING DATE: 2000-03-24

EARLIER APPLICATION NUMBER: 60/125,913

EARLIER FILING DATE: 1999-03-24

NUMBER OF SEQ ID NOS: 26

SOFTWARE: patentin Ver. 2.1

SEQ ID NO 20

LENGTH: 208

TYPE: PRT

ORGANISM: Canis familiaris

US-09-535-521-20

Query Match 17.58: Score 296.5: DB 4: Length 208.

Best Local Similarity 31.58: Pred. No. 8.5e-20:

Matches 68: Conservative 42: Mismatches 81: Indels 72

QY 123 EKSKLOEYQELTRKAAVGLPEKSKLOEYQELTRKAAVGLPEKSKLTG 78

Db 1 QKSGAAVSVQMKREIQAQKRM--KAQSELSQNLIALPSLNNKLSQSLNI

QY 181 TDLTAERL-----CRHCQKWTFFQGNFYFMSNSQRNMHDSVTA 128

Db 59 ERDQEVYKLMELHVSNGSEQNTQPKMLNFOKQVYFGDEPKKMLDAPFA

QY 230 VVIRAFQNTQQTQSRNFSNMRI SQNGESTWAWVGSPLSPFYKWS

Db 119 ASHSOEQDLARYANKG--TWIGLQDLDRSEFLIMLENPLNS--NK

QY 290 G-NEKAFSSSG-WNDNRG--IVINWYVCKKFAAG 321

Db 174 GGGFTVWMCSSQWMDAFQSSSLQGWTPPLATG 208

RESULT 15

5514582-11

Patent No. 5514582

APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.

TITLE OF INVENTION: RECOMBINANT JMA ENVYING HYBRID

IMMUNOGLOBULINS

NUMBER OF SEQUENCES: 43

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 05/05/185,670

FILING DATE: 21-JAN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 986,941

FILING DATE: 08-DEC-1992

APPLICATION NUMBER: 808,122

FILING DATE: 16-DEC-1991

APPLICATION NUMBER: 440,625

FILING DATE: 22-NOV-1989

APPLICATION NUMBER: 315,015

FILING DATE: 23-FEB-1989

SEQ ID NO. 11.

LENGTH: 111

5514582-11

GenCore version 5.1.4
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 11:20:21 Search time: 66 seconds

(without alignments)
1779 820 Million coll updates/sec

Title: US-09-831-458A-12

Perfect score: 1696

Sequence: 1 MDSKEPPVQQLITLCTGH

Scoring table: RIGSUM62

Xgapop 10.0, Xgapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-O=/cgn2_1/705PT0_SF001/US09831458/rnall-05122002_10x16_1516x73pf-pretty.fasta_1.519
-DB-Issued_Patents_NA -OFMT-Istlap -SUFIX-std.rni -MINMATCH-0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS-bits -START-1 -END-1 -MATRIX-bioSUM62 -TRANS-human40.cdi
-LIST-45 -DOCALLIGN-200 -THR-SCORE-pct -THR-MAX-100 -THR-MIN=0 -ALIGN-15
-MODE-LOCAL -OUTFMT-p2n -NORM-ext -HEAVISIZE=500 -MINLEN=0 -MAXLEN=200000000
-USFR-US09831458.acn_1.1.17-strucall-05122002_103216_15169 NCTU-6 -ICPU-3
-NO_XLXAY -NO_MMAP -LARGEQUERY -NPG_SCORES=0 -WAIT-1 -ONCOL=0 -REV-TIMEOUT=120
-WARN-TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database: Issued_Patents_NA:*

1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*

2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*

3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*

4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*

5: /cgn2_6/prodata/1/ina/6C.COMB.seq:*

6: /cgn2_6/prodata/1/ina/6D.COMB.seq:*

7: /cgn2_6/prodata/1/ina/6E.COMB.seq:*

8: /cgn2_6/prodata/1/ina/6F.COMB.seq:*

9: /cgn2_6/prodata/1/ina/6G.COMB.seq:*

10: /cgn2_6/prodata/1/ina/6H.COMB.seq:*

11: /cgn2_6/prodata/1/ina/6I.COMB.seq:*

12: /cgn2_6/prodata/1/ina/6J.COMB.seq:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1412.5	83.3	1212	4	US-09-591-435-9
2	1412.5	83.3	1312	4	US-09-517-605-1
3	1407	83.0	1643	4	US-09-517-605-16
4	1398.5	82.5	1212	4	US-09-591-435-10
5	1393.5	82.2	1312	4	US-09-591-435-11
6	3350	20.6	1005	1	US-08-365-1038-1
7	338	20.0	924	1	US-08-365-1038-5
8	335	19.8	885	1	US-08-365-1038-3
9	329.5	19.4	876	4	US-09-535-521-4
10	329.5	19.4	876	4	US-09-535-521-6
11	329.5	19.4	2851	4	US-09-535-521-1
12	329.5	19.4	2851	4	US-09-535-521-3

13	322.5	19.0	1047	1	US-08-465-1038-7
14	319.5	18.8	1075	1	US-08-465-1038-9
15	319.5	18.8	1458	4	US-09-111-470-3
16	305	18.0	1370	4	US-09-111-470-4
17	296.5	17.5	624	4	US-09-535-521-19
18	296.5	17.5	624	4	US-09-535-521-21
19	292.5	16.7	561	4	US-09-535-521-16
20	292.5	16.7	561	4	US-09-535-521-18
21	263	15.5	417	4	US-09-535-521-10
22	263	15.5	417	4	US-09-535-521-12
23	263	15.5	423	4	US-09-535-521-13
24	263	15.5	423	4	US-09-535-521-15
25	256	15.1	384	4	US-09-535-521-7
26	256	15.1	384	4	US-09-535-521-9
27	246	14.5	1318	2	US-08-809-494A-5
28	246	14.5	1318	2	US-08-809-494A-7
29	243.5	14.4	369	4	US-09-535-521-24
30	243.5	14.4	369	4	US-09-535-521-26
31	232	13.7	1104	4	US-09-111-470-1
32	230	13.6	501	4	US-08-772-440-20
33	230	13.6	1227	4	US-08-772-440-20
34	228	13.4	393	4	US-08-772-440-20
35	226.5	13.4	1418	4	US-09-111-470-7
36	225.5	13.3	1897	2	US-08-809-494A-1
37	225.5	13.3	1897	2	US-09-452-402-1
38	225.5	13.3	1906	2	US-08-809-494A-3
39	225.5	13.3	1906	4	US-09-452-402-3
40	216.5	12.8	402	4	US-08-543-246B-10
41	216.5	12.8	648	4	US-08-543-246B-14
42	216.5	12.8	1755	4	US-08-543-246B-8
43	202.5	11.9	4771	4	US-08-840-062-1
44	200.5	11.8	4588	4	US-08-840-062-1
45	196.5	11.6	871	1	US-08-650-578-1

ALIGNMENTS

RESULT 1
US-09-591-435-9
Sequence 3, Application US/09591435
Patent No. 6280953
GENERAL INFORMATION:
APPLICANT: MESSEIER, WALTER
TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AN
TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED W
FILE REFERENCE: GENO. 200.2
CURRENT APPLICATION NUMBER: 05/07591,435
PRIOR APPLICATION NUMBER: 2000-06-09
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 09/591,435
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 09/240,915
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/074,263
PRIOR FILING DATE: 1998-01-10
PRIOR APPLICATION NUMBER: 60/094,987
PRIOR FILING DATE: 1998-09-02
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 1212
TYPE: DNA
ORGANISM: Homo sapiens
US-09-591-435-9
Alignment scores:
Pred. No.: 7.3e+145
Score: 1412.50
Percent Similarity: 77.04
Best Local Similarity: 71.948
Query Match: 83.288
DB: 4
Length: 1212
Matches: 282
Conservative: 20
Mismatch: 19
Indels: 71
Gaps: 4

[illegible]

Fixed Moves:	7,132	58	Local:	92
Score:	338.50	87	Matches:	92
Percent Similarity:	48.30%		Conservative:	62
Best local Similarity:	28.25%		Mismatches:	19
Query Match:	19.96%		Indels:	41
PR:	1		Gaps:	11


```

QY 46 LysValProSerSerLeuSerGlnGlnGlnSerGlnGlnAspAlaIleIleYr----- 62
Db 3 AAGATCGCAACAGACTGGATATGAGAAAGAAATGAAATGACAAATGAGGATGAG 62
QY 63 -----GlnAsnLeuThrGlnLeu-----Iys 69
Db 63 CTGGCGCTCTCTTTTCTGTCGACTGGAAACGGAAGACATCTAAACACCTGGACAG 122
QY 70 AlaAlaValGlyLeuSerGlnIleYrSerLeuGlnIleIleIleIleIleIleThr 89
Db 123 ACTGCAATGCAATGCTCTGATGCTGATGATGATGATGATGATGATGATGATGATG 176
QY 90 GlnIleuLysAlaAlaValGlyLeuSerGlnIleYrSerLeuGlnIleIleIle 109
Db 177 -----AAATCAATGCTGCTGATGATGATGATGATGATGATGATGATGATG 218
QY 110 GlnLeuThrArgLeuLysAlaAlaValGlyLeuSerGlnIleYrSerLeuGlnIle 129
Db 219 AACTGCAACAGACTGGACATGATGATGATGATGATGATGATGATGATGATGATG 272
QY 130 IleTyrGlnIleuThrArgLeuLysAlaAlaValGlyLeuSerGlnIleYrSer 143
Db 273 CTCTCCCAAGACTGGACATGATGATGATGATGATGATGATGATGATGATGATG 326
QY 150 LeuGlnIleuIleTyrGlnIleuThrArgLeuLysAlaAlaValGlyLeuSer 169
Db 427 AACTCAAAACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 386
QY 170 -----GlnSerLysGlnIleIleIle 176
Db 487 CTGGGCTTGATGAGAGCGGACAGCGCTGGATGCTGATGATGATGATGATGATGATG 446
QY 177 TyrGlnIleuThrArgLeuLysAlaAlaValGlyLeuSerGlnIleYrSerLeu 196
Db 447 GCAAGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 506
QY 197 AspThrPhePheGlnIleLysGlnIleYrSerLeuSerGlnIleYrSerLeu 216
Db 507 AACTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 566
QY 217 AspSerValThrAlaGlyGlnIleValArgAlaGlnLeuValIleIleYrThrAlaGln 236
Db 567 CAGGCGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 626
QY 237 GlnIleuAsnPheLeuGlnIleuThrArgSerAspAlaGlyPheSerIlePheLeu 256
Db 627 GAACAGGACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 680
QY 257 SerAspLeuAsnGlnIleGlyThrIlePheIleValAspIleYrSerProLeuSerProSer 276
Db 681 CAGGATCTCAATATGAGGAGGAGATGATGATGATGATGATGATGATGATGATGATG 740
QY 277 PheIleuArgIleTyrPheAsnSerGlyGlnProAsnAsnSerGly---AsnGlnAspGlyAla 295
Db 741 -----AACTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 791
QY 296 GlnPheSerGlySerGly---TyrPheAspAsnArgGly---AspValAspAsnIleTyrPhe 313
Db 792 ATGATGCGGAGATGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 851
QY 314 IleGlyLysIleGlyProAlaIleGly 321
Db 852 GTGTGTGAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 875

```

```

TITLE OF INVENTION: DNA Sequences for Soluble Forms of
TITLE OF INVENTION: and Methods of Use for Same
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESS: Zartley, McKee, Thumby, Voorhees & Sease
STREET: 901 Grand Ave. Suite 4200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-09/465,103B
FILING DATE: 28-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/WORK NUMBER: 5111 95-24
TELEPHONE: (515) 288-3667
TELEFAX: (515) 288-1338
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 base pairs
TYPE: nucleic acid
STRANDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 24..884
US-09-831-458a-12 (1-325) x US-08-465-103B-3 (1-885)
Alignment Scores:
Pred. No.: 1,616 27 Length: 885
Score: 335.00 Matches: 94
Percent Similarity: 48.21% Conservative: 57
Best Local Similarity: 29.64% Mismatches: 99
Query Match: 19.75% Indels: 60
DB: 1 Gaps: 14
QY 46 LysValProSerSerLeuSerGlnGlnGlnSerGlnGlnAspAlaIleIleYr----- 62
Db 3 AAGATCGCAACAGACTGGATATGAGAAAGAAATGAAATGACAAATGAGGATGAG 62
QY 59 -----AspAlaIleIleYrGlnAsnLeuThrGlnLeu----- 69
Db 63 AATCTAAACAGCTGGAGAGCTGCAATTCATTAATGATGCTCTGATCTATA 740
QY 69 ---LysAlaAlaValGlyLeuSerGlnIleYrSerLeuGlnIleIleIleIleIle 109
Db 123 CAAAATTCAGAGTAAATCAATTCAGAGTAAATTCAGAGTAAATTCAGAGTAAATTC 176
QY 188 LeuThrGlnLeuLysAlaAlaValGlyLeuSerGlnIleYrSerLeuGlnIleIle 218
Db 183 TCCGAGAACCTGACCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 272
QY 108 TyrGlnIleuThrArgLeuLysAlaAlaValGlyLeuSerGlnIleYrSerLeu 143
Db 237 TCCGAGAACCTGACCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 272
QY 128 GlnIleuLysIleGlnIleuThrArgLeuLys----- 169
Db 291 TCAAAACTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 326

```


GenCore version 5.1.4
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_f2n model

Run on: December 7, 2002, 11:18:56, Search time 61 Seconds
(without alignments)
2077,510 Million cell updates/sec

Title: US-09-831-458A-12

Percent score: 1696
1 MCODESPRVGGTGTTCATGSH PRTVNTWVTKRPACPPPP 325

Scoring table:

BLISUM62	
Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700950

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_model -DEV=xlh
-O/cg02_1/US09-831-458A-12/002_109217_15180/3ff_query_fast4_1_519
-DB=Published.Applications.NA -QMT=fastap -SUFFIX=std rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blisum62
-TRANS=human40.cdi -LIST=45 -DOCFALIGN=200 -THR_SCORE=ext -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0
-MATEL=2000000000 -USFR=US09831458 -CGN=1_1_21 -rnat_05122002_109217_15180
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAP -IARROWDIR= -NPG_SCORES=0 -WAIT -TONTLOG
-DEV_TIME=007-120 -WARN_TIME=007-10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

1: /cg02_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cg02_6/prodata/2/pubpna/pcr_NEM_PUB.seq.*
3: /cg02_6/prodata/2/pubpna/US06_NEM_PUB.seq.*
4: /cg02_6/prodata/2/pubpna/US06_NEM_PUB.seq.*
5: /cg02_6/prodata/2/pubpna/US07_NEM_PUB.seq.*
6: /cg02_6/prodata/2/pubpna/US07_NEM_PUB.seq.*
7: /cg02_6/prodata/2/pubpna/US08_NEM_PUB.seq.*
8: /cg02_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cg02_6/prodata/2/pubpna/US09_NEM_PUB.seq.*
10: /cg02_6/prodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cg02_6/prodata/2/pubpna/US10_NEM_PUB.seq.*
12: /cg02_6/prodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cg02_6/prodata/2/pubpna/US10_NEM_PUB.seq.*
14: /cg02_6/prodata/2/pubpna/US10_PUBCOMB.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	329	19.4	1617	US-09-728-952-45	Sequence 45, Appl
2	19.5	18.8	1458	US-09-862-802-3	Sequence 3, Appl
3	305	18.0	1370	US-09-862-802-9	Sequence 9, Appl
4	296.5	17.5	1355	US-09-978-657-230	Sequence 230, App

Result No.	Score	Query Match	Length	DB ID	Description
5	296.5	17.5	1355	US-09-978-657-230	Sequence 230, App
6	296.5	17.5	1355	US-09-978-657-230	Sequence 230, App
7	289	17.0	592	US-09-864-761-7450	Sequence 7450, App
8	284	16.7	152	US-09-864-761-24072	Sequence 24072, App
9	275.5	16.2	1277	US-09-880-107-2230	Sequence 2230, App
10	274.5	16.2	2099	US-09-905-291A-376	Sequence 376, App
11	273.5	16.1	997	US-09-909-330-376	Sequence 376, App
12	273.5	16.1	997	US-09-909-330-376	Sequence 376, App
13	273.5	16.1	997	US-09-909-330-376	Sequence 376, App
14	273.5	16.1	997	US-09-909-330-376	Sequence 376, App
15	267.5	15.8	1300	US-09-880-107-2731	Sequence 2731, App
16	267	15.7	1309	US-09-880-107-2731	Sequence 2731, App
17	263.5	15.5	549	US-09-880-107-2731	Sequence 2731, App
18	263.5	15.5	549	US-09-880-107-2731	Sequence 2731, App
19	263	15.5	3750	US-09-917-800A-474	Sequence 474, App
20	253.5	14.9	1036	US-09-843-381-119	Sequence 119, App
21	249.5	14.7	758	US-09-843-381-119	Sequence 119, App
22	246.5	14.5	1290	US-09-917-800A-1642	Sequence 1642, App
23	244.5	14.3	3763	US-09-870-759-141	Sequence 141, App
24	232	13.7	1091	US-09-764-870-199	Sequence 199, App
25	232	13.7	1091	US-09-764-870-199	Sequence 199, App
26	232	13.7	1104	US-09-962-802-1	Sequence 1, App
27	230	13.6	2930	US-09-745-763-198	Sequence 198, App
28	226.5	13.4	1418	US-09-862-802-7	Sequence 7, App
29	225	13.3	1209	US-09-052-586-593	Sequence 593, App
30	215	12.7	693	US-09-843-381-994	Sequence 994, App
31	211.5	12.5	443	US-09-960-352-7063	Sequence 7063, App
32	210.5	12.4	5185	US-09-870-759-94	Sequence 94, App
33	206	12.1	759	US-09-944-807-5	Sequence 5, App
34	202.5	11.9	562	US-09-744-870-124	Sequence 124, App
35	202.5	11.9	1066	US-09-744-870-15	Sequence 15, App
36	201	11.9	528	US-09-843-381-995	Sequence 995, App
37	201	11.9	743	US-09-764-870-16	Sequence 16, App
38	195	11.5	1238	US-09-992-598-356	Sequence 356, App
39	195	11.5	1238	US-09-992-598-356	Sequence 356, App
40	195	11.5	1238	US-09-989-727-356	Sequence 356, App
41	195	11.5	1238	US-09-989-727-356	Sequence 356, App
42	195	11.5	1238	US-09-989-727-356	Sequence 356, App
43	195	11.5	1238	US-09-989-727-356	Sequence 356, App
44	195	11.5	1238	US-09-989-727-356	Sequence 356, App
45	195	11.5	1238	US-09-989-727-356	Sequence 356, App

ALIGNMENTS

RESULT 1
US-09-728-952-45
Sequence 45, Application US/09728952
Parent No. US2002011102A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Wang, Jian-Fui
APPLICANT: Wang, Dunru
APPLICANT: Yamazaki, Yuki
APPLICANT: Ujwal, Manusha L.
APPLICANT: Drmanac, Radovic L.
TITLE OF INVENTION: No. US2002011102A1 Nucleic Acids
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 799
CURRENT FILING DATE: 2000-11-30
NUMBER OF SEQ ID NOS: 101
SOFTWARE: pL_genes Version 2.0
SEQ ID NO 45
LENGTH: 1617
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS


```

QY 121 SerGlyGlnGlnIleTyrGlnIleThrAspLeuLysThrAlaPheGluArgLeu 160
DB 776 GCC-----TCCACTGGAAGGACG 793
QY 191 CysAlaGHisCysProLysAspTrpThrPheThiGlnLysGlyStryrThiMetGlyAsn 210
DB 744 TGC-----TCCCGGCGTAACTGGTGGATGATTAAGACACCTGCTACTGTTCTTCTTCCAC 847
QY 211 SerGlnAsnTrpHisLaspSerValThrAlaCysGlnGlnIleValaLacGlnIleVal 230
DB 848 TCTGGAGTCTCTCTGGCGAGAGAGCTGACAGACTACTGCTACCTGAAGAACAGCGGACCTGCTG 907
QY 231 ValIleLysThrAlaGlnIleGlnIleAsnPheLeuGlnIleLeuGlnIleSerAlaSerAspVal 250
DB 908 GTGATCACTGATGAGGAGAGCAATTTGTGTGCAAAATATCTA-----GGTCTGTCTGA 961
QY 251 PheSerTrpMetGlyLeuSerAspLeuAsnGlnGlnIleGlyTrpTrpGlnTrpValAspGly 270
DB 962 TACACTGATGATGAGCTGATAGTGAAC-----CTTAACTAACTATGAAATGAGATGAGATG 1015
QY 271 SerProLeuSerProSerPheGluArgTyrTrpAsnSerGlyIleProAsnAsn----- 288
DB 1016 ACAGACTATGGAGCGCGCTTCTTACAGAC-----TGGAAAGGTAAGTCTCAATCAATCAATCTG 1072
QY 289 -----SerGlyAsnGlnAspCysAlaGlnIlePheSerGlyStryr---TrpAsn 303
DB 1073 GAGATATGAACTAGGTTGATGAGAGAGAGTGGTGGTGAATCTTAACTGATGATGAGAGAG 1132
QY 304 AspAsnArgCysAspValAspAsnTyrTrpLeuCysLys 316
DB 1133 GACGAGCTGTCTGCAAGAGCTTACACTGATGATGATGAGAG 1171

RESULT 3
US-09-862-802-9
: Sequence 9, Application US/0986-802
: Patent No. US20020165346A1
: GENERAL INFORMATION:
: APPLICANT: Schering-Plough Corporation
: TITLE OF INVENTION: MAMMALIAN MEMBRANE PROTEIN GENES; RELATED REAGENTS
: FILE REFERENCE: SP06958
: CURRENT APPLICATION NUMBER: US/09/862-802
: PRIOR FILING DATE: 2002-01-10
: PRIOR APPLICATION NUMBER: US 09/111,470
: PRIOR FILING DATE: 1998-07-08
: NUMBER OF SEQ. ID NOS: 11
: SOFTWARE: PatentIn version 3.1
: SEQ. ID NO 9
: LENGTH: 1370
: TYPE: DNA
: ORGANISM: unknown
: FEATURE:
: OTHER INFORMATION: mammalian nucleic acid
: NAME/KEY: CDS
: LOCATION: 273...1091
: OTHER INFORMATION: Protein coding sequence
US-09-862-802-9

Alignment Scores:
Pred. No.: 3,530-24 Length: 1370
Score: 305.00 Matches: 89
Percent Similarity: 45.34% Conservatio: 49
Best Local Similarity: 29.28% Mismatches: 96
Query Match: 17.98% Indels: 70
DB: 9 Gaps: 15

US-09-841-458A-12 (1-325) x US-09-862-802-9 (1-1370)
QY 20 HisGlyAlaLeuValLeuGlnIleLeuSerPheMetLeuLeuAlaGlyValLeuValAla 39
DB 346 AATGGGCACTTCTCTCCATCCATCCCTC-----CTGCGCTGGTGGTCAATCATATGCTG 386
QY 40 IleLeuValGlnValSerLysValProSerSerLeuSerGlnGlnIleGlnIleGlnIleAsp 59

```

```

DB 387 GTTGATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCG 441
QY 60 AlaIleTyrGlnAsnLeuThrIleValGlnAlaValGlyGluLeuSer 441
DB 441 -----TTTACCACTTCAC----- 462
QY 80 LysLeuGlnGlnIleTyrGlnIleLeuThrGlnIleGlnIleGlnIleGlnIleGlnIle 462
DB 462 ACTGTGGCGGAGATC-----CAGGATGATCTC----- 483
QY 100 GlnLysSerLysLeuGlnIleIleTyrGlnIleLeuThrArgLeuLysAla 483
DB 483 -----TTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 504
QY 120 GlnIlePheGlnLysSerLysLeuGlnIleIleTyrGlnIleLeuThrArgLeuLysAla 504
DB 504 GTTTC-----AAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 525
QY 140 AlaValGlyGlnIleProGlnLysAlaLysLeuGlnIleIleTyrGlnIle 525
DB 525 -----TTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
QY 160 LeuGlyAlaValAlaValGlyGlnIleProGlnLysAlaLysLeuGlnIleIleTyrGlnIle 546
DB 546 -----TTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 567
QY 180 LeuThrAspLeuLysThrAlaPheGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 567
DB 567 -----GCTTCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 588
QY 200 PheThrGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 588
DB 588 -----TTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609
QY 220 ThrAlaCysGlnGlnIleValAlaValGlyGlnIleProGlnLysAlaLysLeuGlnIleIleTyrGlnIle 609
DB 609 -----TTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 630
QY 240 PheLeuGlnIleLeuGlnIleThrSerArgSerAsnArgPheSerTrpMetGlyLys 630
DB 630 -----TTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
QY 260 AsnGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 651
DB 651 -----TTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672
QY 280 TyrTrpAsnSerGlyGlnIleProGlnLysAlaLysLeuGlnIleIleTyrGlnIle 672
DB 672 -----TTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 693
QY 294 CysAlaGlnIlePheSerLysLeuGlnIleIleTyrGlnIleLeuThrArgLeuLysAla 693
DB 693 -----TTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 714
QY 947 TGTCTCTCACTTCCATCCAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1047
QY 313 TrpIleCysLys 316
DB 1047 TGGTGTCTGAG 1058

RESULT 4
US-09-978-295A-230
: Sequence 230, Application US/09978-295A
: Patent No. US2002015006A1
: GENERAL INFORMATION:
: APPLICANT: Ashkenazi, Avi
: APPLICANT: Baker, Kevin P.
: APPLICANT: Bolstein, David
: APPLICANT: Desnoyers, Luc
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Fong, Sherman
: APPLICANT: Gao, Wei-Oliang

```

APPLICANT: Gerber, Hanspeter
APPLICANT: Gottisen, Mary F.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kiljavan, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James,
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Thomas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William L.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C11
CURRENT APPLICATION NUMBER: US/09/978,295A
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/9418545
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/004249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/005411
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/06364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/07632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078846
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079456
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079554
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082894
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/082336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083600
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742

APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT APPLICATION NUMBER: US/09/4978,647
PRIOR APPLICATION NUMBER: 2001-19-17
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/948585
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078946
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078940
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078949
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-04-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079693
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333

PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081917
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081919
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081938
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083346
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083342
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083445
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07

[illegible]

```

: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO: 2230
: LENGTH: 1277
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. U520020142081A1 M10953
US-09-831-458a-12 (1-325) x US-09-880-107-2230 (1-1277)

```

```

Alignment Scores:
Pred. No.: 5,31e-21 Length: 1277
Score: 275.50 Matches: 91
Percent Similarity: 39.60% Conservative: 46
Best Local Similarity: 26.30% Mismatches: 114
Query Match: 16.24% Indels: 95
DB: 10 Gaps: 14

```

```

US-09-831-458a-12 (1-325) x US-09-880-107-2230 (1-1277)
QY 6 GlnProArgValGlnGlnLeuGlyLeuLeuGlyLeuLeuGlyLeuValLeu 25
DB 257 CAGGCTCTCTGAGAGCTCTG-----TCTCTGAGCTCTGCTCTCTCTCT 304
QY 26 GlnLeuLeuSerPheMetLeuLeuAlaGlyValLeuValAlaLeuValSer 45
DB 305 TCCTCTG-----GCTCTGAGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 343
QY 46 LysValIleProSerSerLeuSerGlnGlnGlnSerGlnAlaIleTyrGln 65
DB 344 GTGATCGATCCCAAACTCCCACTCTCAAGAGAGCTCTGAGAGAGAGAGAGAG 403
QY 66 ThrGlnLeuLysAlaValAlaGlyLeuSerGlyLeuSerGlyLeuGlnIleTyr 85
DB 404 AGCAACTTCACAGCGACGACG-----GAC 427
QY 86 GlnGlnLeuThrGlnLeuLysAlaValAlaGlyLeuProGlyLeuSerLysLeuGln 105
DB 428 GCCCAGGCTCAAGGCTTGAACACCGACCGAGAGATGCGAAGAAAGAGAGTGGTA 487
QY 106 GlnIleTyrGlnGlnIleThrArgLeuLysAlaValAlaGlyLeuProGlyLeuSer 125
DB 488 GAG-----TCT 493
QY 126 LysLeuGlnGlnIleTyrGlnGlnLeuThrArgLeuLysAlaValAlaGlyLeuPro 145
DB 494 CAGCTGAGAGAAATAGTAAAGAAATTCAGTCAAA----- 526
QY 146 GlnIleSerLysGlnGlnGlnIleTyrGlnGlnLeuThrArgLeuLysAlaValAlaGly 165
DB 527 GATCACTCCAGCCG-----CTGCTCCAGCTCAAGGAGCTGGTCT 568
QY 166 GlnLeuProAspGlnSerLysGlnGlnGlnIleTyrGlnGlnLeuThrAspLeuGln 185
DB 569 GACTGTGGAGACTGAGCTCTTCAGATGAGGCGCGCTGAG-----GACAA 613
QY 186 AlaPheGlnArgLeuGlySerGlnGlnGlnSerLysAspTyrThrPhePheGlnIleTyrGln 205
DB 614 GAGCTCAAAAAGAGAGCTG-----TCCCGCTCAAGCTCTGAGTCAAGCGCAAGTCT 667
QY 206 TyrPheMetSerAspSerGlnArgAspTyrPhePhePheGlnIleVal 225
DB 668 TATGGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 727
QY 226 ArgAlaGlnLeuValAlaIleTyrThrAlaGlnGlnGlnGlnGlnGlnGlnGln 245
DB 728 GAGCGTCACTGAGTGTGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 787

```

```

QY 246 SerArgSerAsnArgPheSerTrpMetIleLeuSerAspLeuAsnIleVal 265
DB 788 GCGCCCTGTGAGC-----ACCTGATGAGCGCTGAGAGC-----GAAAG 304
QY 266 GlnTyrValAspGlySerProLeuSerTrpSerPheGlnArgTyrTrpAsp 285
DB 836 AAGTAACTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 885
QY 286 TrpAsnAsn-----SerGlnAsnIleAspGlyAlaGln 304
DB 893 CCGGAGCACTGCTAGCGGCAAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAG 343
QY 300 SerGly TrpAspAspAspAspTyrAspValAspAspTyrTrpLeu 325
DB 954 CACGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 364
QY 317 ----- 385
DB 1013 CTGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 424
QY 319 AlaAlaCysPheArgAsp 424
DB 1073 CCGCGAGAGGCTCGCGCAT 1096
RESULT 10
US-09-764-870-14
: Sequence 14, Application US/09764870
: Patent No.: US20020042486A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibody
: FILE REFERENCE: PT214
: CURRENT APPLICATION NUMBER: US/09764,870
: CURRENT FILING DATE: 2001-01-17
: PRIOR APPLICATION DATA REMOVED - CONSULT PALM or FILE WR
: NUMBER OF SEQ ID NOS: 646
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 14
: LENGTH: 2099
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-764-870-14
Alignment Scores:
Pred. No.: 1.39e-20 Length: 2099
Score: 274.50 Matches: 91
Percent Similarity: 44.30% Conservative: 51
Best Local Similarity: 27.18% Mismatches: 124
Query Match: 16.19% Indels: 44
DB: 10 Gaps: 12
US-09-831-458a-12 (1-325) x US-09-764-870-14 (1-2099)
QY 36 ValLeuValAlaIleLeuValGlnValSerIleValProSerSerLeu 385
DB 174 CTGTGTGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 424
QY 56 SerGlnGlnAspAlaIleTyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 445
DB 234 GAT-----GTAAGAGTAAAGTCTGAGAGAGAGAGAGAGAGAGAGAGAG 484
QY 76 SerGlnGlnSerLysLeuGlnGlnIleTyrGlnGlnGlnGlnGlnGlnGln 504
DB 282 AGCAGCCCTGAGT-----TCTCAATTAAGAAAGAAAGTAAAGTAAAGTAAAG 543
QY 96 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 562
DB 336 GTTATGATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 591
QY 115 LysAlaAlaValAlaGlyLeuProGlnIleLysSerLysLeuGln 610
DB 356 AAAGCAGCTGTG-----GACAAAGCTCAAGCTCAAGCTCAAGCTCAAGCT 649

```


PRIOR FILING DATE:	1997-10-31
PRIOR APPLICATION NUMBER:	60/0064103
PRIOR FILING DATE:	1997-10-31
PRIOR APPLICATION NUMBER:	60/0065311
PRIOR FILING DATE:	1997-11-13
PRIOR APPLICATION NUMBER:	60/0066120
PRIOR FILING DATE:	1997-11-21
PRIOR APPLICATION NUMBER:	60/0066466
PRIOR FILING DATE:	1997-11-24
PRIOR APPLICATION NUMBER:	60/0066772
PRIOR FILING DATE:	1997-11-24
PRIOR APPLICATION NUMBER:	60/0069335
PRIOR FILING DATE:	1997-12-11
PRIOR APPLICATION NUMBER:	60/0069425
PRIOR FILING DATE:	1997-12-12
PRIOR APPLICATION NUMBER:	60/0069470
PRIOR FILING DATE:	1997-12-17
PRIOR APPLICATION NUMBER:	60/0068017
PRIOR FILING DATE:	1997-12-18
PRIOR APPLICATION NUMBER:	60/0077450
PRIOR FILING DATE:	1998-03-10
PRIOR APPLICATION NUMBER:	60/0077632
PRIOR FILING DATE:	1998-03-11
PRIOR APPLICATION NUMBER:	60/0077649
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/0079786
PRIOR FILING DATE:	1998-03-27
PRIOR APPLICATION NUMBER:	60/0080107
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/0080194
PRIOR FILING DATE:	1998-03-31
PRIOR APPLICATION NUMBER:	60/0080547
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/0080333
PRIOR FILING DATE:	1998-04-01
PRIOR APPLICATION NUMBER:	60/0080149
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/0081070
PRIOR FILING DATE:	1998-04-08
PRIOR APPLICATION NUMBER:	60/0081195
PRIOR FILING DATE:	1998-04-09
PRIOR APPLICATION NUMBER:	60/0081838
PRIOR FILING DATE:	1998-04-15
PRIOR APPLICATION NUMBER:	60/0082588
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION NUMBER:	60/0082589
PRIOR FILING DATE:	1998-04-21
PRIOR APPLICATION NUMBER:	60/0082704
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/0082797
PRIOR FILING DATE:	1998-04-22
PRIOR APPLICATION NUMBER:	60/0083322
PRIOR FILING DATE:	1998-04-28
PRIOR APPLICATION NUMBER:	60/0083495
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083496
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083499
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0083559
PRIOR FILING DATE:	1998-04-29
PRIOR APPLICATION NUMBER:	60/0084366
PRIOR FILING DATE:	1998-05-05
PRIOR APPLICATION NUMBER:	60/0084414
PRIOR FILING DATE:	1998-05-06
PRIOR APPLICATION NUMBER:	60/0084639
PRIOR FILING DATE:	1998-05-07

[illegible]

```

: PRIOR FILING DATE: 1998-06-12
: PRIOR APPLICATION NUMBER: 60/089512
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089514
: PRIOR FILING DATE: 1998-06-16
: PRIOR APPLICATION NUMBER: 60/089538
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089598
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089653
: PRIOR FILING DATE: 1998-06-17
: PRIOR APPLICATION NUMBER: 60/089608

```

Alignment Scores:

```

Pred. No.: 6,150-21 Length: 997
Score: 273.50 Matches: 60
Percent Similarity: 54.49% Conservative: 25
Best Local Similarity: 38.46% Mismatches: 56
Query Match: 16.13% Indels: 15
DB: 12 Gaps: 6

```

US-09-831-458a-12 (1-325) x US-10-052-586-23 (1-997)

```

QY 172 LysGlnGlnGlnIleTyrglnGlnLeuThaAspLeuLys-----ThraAla 186
DB 268 AAAAAGTTTCACTACAGCAATTTCACAGAGCTCTCGTACATTATGATCAGGT 327
QY 187 PheGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 206
DB 328 TCAGTCAGCAATTGT-----TTCACATTCACTGCAATATTTTCAATCCAGCTCAG 381
QY 207 PheMetSerAsnSerGlnIleAsnSerPheLysSerValThraLysGlnGlnGlnGln 226
DB 382 TTCCTTCTTCACTGACACCAATTCCTGCTGCTTAAAGACAGCTGACAGCAAGTGG 441
QY 227 AlaGlnLeuValValIleTygIleThraGlnGlnGlnGlnGlnGlnGlnGlnGln 246
DB 442 GCTAACCTGGGGTTATCACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
QY 247 AsnSerAsnMetPheSerTyrMetGlyIleGlnSerAspLeuAsnGlnGlnGlnGln 266
DB 502 AAAATGAGAGAGTTT--TTATTGAGCTGTCTGACAGCTGTGAGAGAGAGAGAGAG 558
QY 267 TrpValAspGlnSerProLeuSerProSerPheGlnIleGlnGlnGlnGlnGln 286
DB 559 TGGGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 615
QY 287 AsnAsnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 300
DB 616 AACACATACCTACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 675
QY 301 GlyTrpAsnAspAsnArgGlnAspValAspAsnTyrTrpIleGlnGlnGlnGln 316
DB 676 AATTGATATGATGATTAACCTCTTCTGCTGCAATATTTCGATTTGTCGA 723

```

RESULT 15

```

: Sequence 3731, Application US/06890107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: PRIOR FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950

```

```

: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO: 3731
: LENGTH: 1400
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 X55283
US-09-880-107-3731

```

Alignment Scores:

```

Pred. No.: 4,076-26 Length: 1300
Score: 267.50 Matches: 95
Percent Similarity: 44.31% Conservative: 49
Best Local Similarity: 29.23% Mismatches: 106
Query Match: 15.77% Indels: 76
DB: 10 Gaps: 16

```

US-09-831-458a-12 (1-325) x US-09-880-107-3731 (1-1300)

```

QY 6 GluProArgValIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 25
DB 272 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 316
QY 26 GluLeuLeuSerPheMetLeuLeuAlaGlnValLeuValAlaIleValGlnValSer 45
DB 317 CTGGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 358
QY 46 LysValProSerSerLeuSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 65
DB 359 GTCAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 397
QY 56 PheMetSerAsnSerGlnIleAsnSerPheLysSerValThraLysGlnGlnGlnGln 85
DB 498 GCTAACCTGGGGTTATCACTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
QY 86 GlnGlnLeuThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 105
DB 452 CAGAGTATATATATATATATATATATATATATATATATATATATATATATATAT 487
QY 109 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 125
DB 488 -----AGAGTCAATCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
QY 139 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 145
DB 509 AAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556
QY 140 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 165
DB 557 -----CTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 583
QY 166 GluLeuArgAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 185
DB 584 GAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 628
QY 186 AlaPheGlnIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 205
DB 629 GCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 682
QY 206 TyrPheGlnSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 225
DB 683 TACGATGTTTCTTCACTGACAGCAATTTCACAGAGCTCTCGTACATTATGATCAGGT 742
QY 226 ArgAlaGlnIleValIleTygIleThraGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 245
DB 743 AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 802
QY 246 SerArgSerAsnArgPheSerTyrMetGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 265
DB 803 AACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 850
QY 266 GlnGlnValAspGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 285

```


Genforce version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2002, 09:48:14 : Search time 26 seconds

(without alignments)
518,454 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 1696

Sequence: 1 MSDSKEPRVQGLGLDLSLH.....PCDVNWLCKKPAKPFDE 325

Scoring table: HUSUM62

Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database: SwissProt_40.*

Prod No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	350	20.6	331	1	PCB2_MOUSE
2	326.5	19.3	207	1	LECH_CHICK
3	322.5	19.0	321	1	PCB2_HUMAN
4	309.5	18.2	550	1	KUCR_RAT
5	304	18.2	283	1	LECH_MOUSE
6	307	18.1	304	1	MMGL_MOUSE
7	289	17.0	283	1	LECH_RAT
8	287	16.9	306	1	MMGL_RAT
9	284.5	16.8	548	1	PCB2_MOUSE
10	271.5	16.0	290	1	PCB2_MOUSE
11	267	15.7	311	1	LECH_HUMAN
12	258	15.2	301	1	LECH_MOUSE
13	240	14.2	301	1	LECH_RAT
14	224.5	13.2	216	1	NECD_MOUSE
15	219.5	12.9	162	1	LECH_MOUSE
16	216.5	12.8	216	1	NECD_HUMAN
17	210.5	12.4	146	1	MANE_HUMAN
18	206	12.1	149	1	CLF2_HUMAN
19	202.5	11.9	223	1	NK13_RAT
20	202.5	11.9	2415	1	PCB2_HUMAN
21	201.5	11.9	2124	1	PCB2_MOUSE
22	199	11.7	883	1	PCB2_MOUSE
23	197.5	11.6	179	1	PCB2_MOUSE
24	197	11.6	2132	1	PCB2_MOUSE
25	196.5	11.6	179	1	PCB2_MOUSE
26	195.5	11.5	179	1	CD94_MOUSE
27	195	11.5	883	1	PCB2_MOUSE
28	195	11.5	2109	1	PCB2_MOUSE
29	194.5	11.5	912	1	PCB2_MOUSE
30	194.5	11.5	2333	1	PCB2_MOUSE
31	193	11.4	2363	1	PCB2_MOUSE
32	191	11.4	199	1	PCB2_MOUSE
33	189.5	11.2	1268	1	PCB2_MOUSE

34	188	11.1	227	1	NK11_MOUSE
35	186.5	11.0	1257	1	PCB2_MOUSE
36	180.5	10.6	147	1	LECH_RAT
37	178	10.5	244	1	MANE_MOUSE
38	176.5	10.4	3562	1	PCB2_MOUSE
39	175	10.3	612	1	LECH_MOUSE
40	174	10.3	175	1	ANP_MOUSE
41	171	10.1	223	1	NK12_MOUSE
42	170	10.0	165	1	LECH_RAT
43	169	10.0	125	1	ARAF_MOUSE
44	168	9.9	2738	1	PCB2_MOUSE
45	168	9.9	3358	1	PCB2_MOUSE

ALIGNMENTS

RESULT 1

ID FCB2_MOUSE STANDARD: PRT: **1 AA.

AC FCB2_MOUSE

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Low affinity immunoglobulin epsilon FC receptor (Lymphocyte receptor) (FC-epsilon-RI) (c124).

DE FCB2 OR FCB2A

GN Mus musculus (Mouse).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; Mus musculus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=90017519; PubMed=2529542;

RA Bettler B., Hofstetter H., Kao M., Yokoyama W.M., Kutschera D.H.;

RT "Molecular structure and expression of the murine lymphocyte epsilon receptor for IgE (Fc epsilon RI).";

PL Proc Natl Acad Sci U S A 87:7565-7570(1990).

RN [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=90171598; PubMed=2178457;

RA Gollnick S.O., Trounstein M.L., Yamashita J.C., Kehry M.E., Moore K.W.;

RT "Isolation, characterization, and expression of cDNA encoding the mouse Fc receptor for IgE (Fc epsilon RI).";

RL J. Immunol. 144:1974-1982(1990).

RN [3]

RP SEQUENCE FROM N.A. (FORMS B AND C).

RA MEDLINE=94372613; PubMed 8066226;

RA Kondo H., Ichikawa Y., Nakamura K., Ischiya S.;

RT "Cloning of cDNAs for new subtypes of murine low-affinity Fc epsilon receptor (Fc epsilon RI/C124).";

RL Int. Arch. Allergy Immunol. 105:48-48(1994).

RN [4]

RP 3D-STRUCTURE MODELING OF LECHIN MAIN.

RA MEDLINE=94191542; PubMed=8142907;

RA Padlan E.A., Helm B.A.;

RT "Modeling of the lectin-homology domains of the human and mouse Fc epsilon receptor (Fc epsilon RI/C124).";

RL J. Biol. Chem. 269:11933-11938(1994).

CC FUNCTION: THIS RECEPTOR HAS ESSENTIAL ROLES IN THE REGULATION OF IgE PRODUCTION AND IN THE DIFFERENTIATION OF B CELLS INTO IgE-PRODUCING PLASMA CELLS.

CC A B-CELL-SPECIFIC ANTIGEN.

CC SUBCELLULAR LOCATION: Type II membrane protein.

CC ALTERNATIVE PRODUCTS: 3 isoforms: A (SHOWN HERE), B AND C.

CC INDUCED BY ALTERNATIVE SPLICING.

CC PTM: N- AND O-GLYCOSYLATED (BY SIMILARITY).

CC DIFFER IN BOTH STRUCTURE AND FUNCTION: HIGH AFFINITY FOR IgE (FORM A) AND LOW AFFINITY RECEPTORS (FORMS B AND C).

CC AND MONOCYTES.


```

FT REPEAT 111 131
FT DISULFID 160 288 BY SIMILARITY.
FT DISULFID 163 174 BY SIMILARITY.
FT DISULFID 191 282
FT DISULFID 259 273
FT CARBOHYD 63 63 N-LINKED (GLCNAc, ...) (POTENTIAL).
FT VARSPIC 1 7 MEGAYS -> MPPSO (IN ISOFORM B).
FT CONFLICT 269 269 N -> T (IN REF. 3).
SU SEQUENCE 421 AA: 36468 MW: 48670 RC:06515887 CPr64:

Query Match: 19.0%, Score 122.5, DB 1, Length 321.
Best Local Similarity 28.6%; Pred. No. 9a-14;
Matches 95; Conservative 61; Mismatches 121; Indels 55; Gaps 14;

QY 1 MSDEKPRVQGIIGLGTGHC--IVGLISFMIAGVAVIIVQSKVPSISQ--EQLF 57
DB 1 MEEQVSPFIEELPRRCRCRGTQIVAGIVTALMAGLLILMLMWDITQSLKLEERA 60
QY 58 QDAIYQNIQIKAAGV-ELSEKSKLQEIYQETIOLKAAGELPEKSKLQEIYQETIKA 116
DB 61 ARNVSQVSKNLSHHGDMQKQSTQISOLEELKA-----EQQRKKS 104
QY 117 AVGLPEPKSKLQEIYQETIOLKAAGVLEPKSK-----LQEIYQETIOLKAAGVLEPKSK 172
DB 105 Q--PLELSMNINLQADLSSEKSO--ELNERNESDLLEPREVETKLR----- 149
QY 173 QQAIVQPIITIKIAFERIPEHCPKRWTFPQVNFPMNSQPMHDSVIAQCPVPAIVV 222
DB 150 -----MELQVSSGVNTCPTEPEKMINPEQRKCYFGKQTKQVHAKYACDIMEQLVSI 201
QY 213 FTANEQNTGIGQTSFNSFTWMLISTINEFTWQWVGSTSLSTSTSTSTSTSTSTSTST 256
DB 202 HSEDEQELTRKHSHTD--SNLZIFILFLKEGFIVWDGSHVYS--NMALDEHSEST 256
QY 292 EYCAEFSSQSG-WNDRGQVQ-NYWICKKPAQ 321
DB 257 EECVMAKRSQPMWDAFQPKLCAWCDPIATF 288

RESULT 4
KUPFER_RAT
ID KUPFER_RAT STANDARD: PRT: 550 AA.
AC P10716:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-type lectin 13 (Kupfer cell receptor).
GN CLEC3F3 OR KCLR.
OS Rattus norvegicus (Rat)
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN 11;
RP SEQUENCE FROM N.A. AND SEQUENCE OF 83-104.
RX MEDLINE:88227339; PubMed 2836387.
RA Hoyle G.W., Hill R.L.;
RT "Molecular cloning and sequencing of a cDNA for a carbohydrate
binding receptor unique to rat Kupfer cells."
FT J. Biol. Chem. 263:7487-7492(1998)
RN 12;
RP SEQUENCE FROM N.A.
RX MEDLINE:91107689; PubMed:1846367;
RA Hoyle G.W., Hill R.L.;
RT "Structure of the gene for a carbohydrate-binding receptor unique to
rat Kupfer cells."
FT J. Biol. Chem. 266:1850-1857(1991)
CC -> FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.
CC COULD BE INVOLVED IN ENOCYTOSIS.
CC -> SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
CC -> TISSUE SPECIFICITY: KUPFER CELLS.
CC -> SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/submit/submit.html).
CC or send an email to license@ebi.ac.uk.
CC
CC EMBL: 307374, AAA1472.1;
DB EMBL: Accession: AAA0092.1;
DB PIR: A28166; A28166.
DB PIR: A28166; A28166.
DB HSP: P20694; HLU1.
DB InterPro: IPR01194, Lectin_C.
DB InterPro: IPR00017; Synaptin.
DB Pfam: PF00059; Lectin_C_1.
DB SMART: SM00034; CLECR_1.
DB SMART: SM00504; SYN_1.
DB PROSITE: PS00615; C-TYPE_LECTIN_1;
DB PROSITE: PS00413; C-TYPE_LECTIN_2;
KW Procyonids; Transmembrane; Glycoprotein; Lectin; Signal-anchor;
KW Endocytosis.
FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 70 550 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 438 538 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 440 536 BY SIMILARITY.
FT DISULFID 516 528 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAc, ...) (POTENTIAL).
FT CARBOHYD 93 93 N-LINKED (GLCNAc, ...) (POTENTIAL).
FT CARBOHYD 145 115 N-LINKED (GLCNAc, ...) (POTENTIAL).
FT CARBOHYD 132 132 N-LINKED (GLCNAc, ...) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAc, ...) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAc, ...) (POTENTIAL).
SU SEQUENCE 550 AA: 61064 MW: 93584 CPr64:306270 CPr64;

Query Match: 18.2%, Score 309.5, DB 1, Length 550;
Best Local Similarity 28.7%; Pred. No. 11e 12;
Matches 81; Conservative 46; Mismatches 110; Indels 45; Gaps 11;

QY 52 SQDESEDAIYQNIQIKAAGVLESEKSKLQEIYQETIOLKAAGVLEPKSK-LQEIYQ 110
DB 286 AQIQNANSHLEFQITQIGIKAKIKSTSTN-----SOLEVYVNRKLKQSSPELIQTRD 339
QY 111 LITFAAVTEPEFESKSKLYGHILKKAVDILPE-----SKLQEIYQETIOLKA 161
DB 340 LSNVSLA-----KSNVQMSQNIQAKAAYQSLKIGLEATIKLAKIQGQSDLEALQK 393
QY 162 AAVGFTIPEQSKQQIYQETITITATAPFELCHRPKRWTFPQVNFPMNSQPMHDSVIA 221
DB 394 AVAAHVGQKQVQVQIQLM-----QWIKYFNKRYFYFSDKKSWEHLENF 439
QY 222 QGVRAQIVITIKIAFERIPEHCPKRWTFPQVNFPMNSQPMHDSVIAQCPVPAIVV 280
DB 440 QVSGAMIAVITIKIAFERIPEHCPKRWTFPQVNFPMNSQPMHDSVIAQCPVPAIVV 497
QY 281 WNSCHTNG SCN EDCAHFGSCGVRRCVWYWIQK 317
DB 438 WESCHTNGMH IHEHEDCVHCGM KWMADIAVYWICK 548

RESULT 5
LECH_MOUSE
ID LECH_MOUSE STANDARD: PRT: 283 AA.
AC P34927; Q64363;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUN-1994 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Asialoglycoprotein receptor 1 (hepatic lectin 1) (MH-1) (ASGP-R)
DE (ASGP-R).
GN ASGRI OR ASGR-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



```

06 150 LGEIYGEITPKAAVGLPFGSKQGVYQETIDKAEFFPLPFGKQWTFQGNVFMS 209
07 121 L-----LLHWKQIVSDVSLSCQMAALR-----GNSERIC--CPINWVEEGSDWFS 167
08 210 NSQRMHNSVACQGVRAQLVAVITAEQNFLOLQTSNFSNFSMGI SLDNFPFTWVWD 266
09 168 NSVKVFTADKYCQLENAHLYVTSWEPFVQGMHPLN--TWIGLTD--QNGPMHWVD 223
10 270 GSPISFQRYWNSGEPNN-----SGNEDEAFSGSGS--NMNMGCDVUNWICK 306
11 224 GIDVETGKKN WRPGQPDWYGHGIGGEGPFAHTTTPYHNMWDVCPPPRYWVE 276

RESULT 8
MMGL_RAT
10 MMGL_RAT STANDARD: PRT: 306 AA.
AC P49301:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (Macrophage
DE galactose/N-acetylgalactosamine-specific lectin) (MMGL).
GN MGL.
OS Rattus norvegicus (Rat)
OC Eukaryota; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE:96293078; PubMed:2358462;
RX Li M., Kurata H., Itoh N., Yamashina I., Kawasaki T.;
RT "Structural similarity between the macrophage lectin specific for
RT galactose/N-acetylgalactosamine and the hepatic asialoglycoprotein
RT binding protein."
RL Biochem Biophys Res Commun. 195:720-725(1998).
CC -!- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYL GALACTOSAMINE
CC UNITS.
CC -!- SUBUNIT: Homo-oligomer.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, contribution
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: J05495; AAA1216.1;
DR HSSP: P06734; IHL;
DR InterPro: IPR001304; Lectin_C;
DR Pfam: PF00059; Lectin_C;
DR SMART: SM00034; CLECT; 1;
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1;
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1;
KW Lectin; Glycoprotein; Transmembrane; Calcium; Signal-anchor.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 58 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 59 306 (POTENTIAL).
FT DOMAIN 174 300 EXTRACELLULAR (POTENTIAL).
FT DISULFID 175 186 C-TYPE LECTIN (LONG FORM).
FT DISULFID 203 298 BY SIMILARITY.
FT DISULFID 276 290 BY SIMILARITY.

```

```

FT CARBOHYD 76 76 N-LINKED (GLYCANC...) (P1)
FT CARBOHYD 168 168 2-LINKED (GLYCANC...) (P1)
SQ SEQUENCE 306 AA: 34242 MW: 34464.966-PHE13 CR 1044
Query Match 16.9% Score: 297.146 1; Length:
Best Local Similarity: 28.0%; Prob: 1.56e-11;
Matches: 86; Conserved: 56; Mismatches: 111; 100%
06 27 LLSFMLAGVAVIIVGSKVPSLSQSGSQAIAVQNLQKAAVDEHSL
07 38 LLLFSL--GLSLLILVLSVLSQNSQLRRLRLTLIDNTS-----N
08 87 ELTQKAAGVLPKSKQIETVGLTRKAAGVLPKSKQIETVGLTR
09 91 RNSLSQTGINSL--KVFEDHQLF---LQAGQ-----LSQVAST
10 147 KSKLQETVGLTRKAAGVLPKSKQIETVGLTRKAAGVLPKSK
11 132 KKE-----QLTRDLSEITD--KVDQLKDKTLTQGLASLKNQSAV
12 201 PLSNLYFMSNSYPMNHISVLAQVRAQVYKIAEPNFIQLOTSNRE
13 182 HEGSCVWFSQSKPWPEDKCYQLNSMLVYVNSLAPNF--LQTHMSV
14 261 QCTQWVGVGSLSTSPQRYKSGERTN-----SGNECAFSGS--WN
15 238 QNSPMRWVGIQYENGF--THNAPKQPMWYGHGIGGEGPFAHTTTP
16 314 TCKRPAA 320
17 297 VCEMKLA 303
DB 297 VCEMKLA 303

RESULT 9
KUCR_MOUSE
10 KUCR_MOUSE STANDARD: PRT: 148 AA.
AC P70194:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C-type lectin 13 (Nupfer cell receptor).
GN CLECSF13 OR KCLR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Granata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BALB/c; Tissue-Liver;
RX Takezawa R., Watanabe H., Nemoto C., Watanabe Y., Akai R.
RT Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AN
CC -!- COULD BE INVOLVED IN ENDOCYTOSIS.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein
CC -!- TISSUE SPECIFICITY: KUPFER CELLS.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN
CC -----
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, contribution
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: D88577; IAAI1647.1;
DR HSSP: P20693; IHL;
DR MGD: MGI:1859834; Clec5f13;
DR InterPro: IPR001304; Lectin_C;
DR Pfam: PF00059; Lectin_C; 1;
DR SMART: SM00034; CLECT; 1;
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1;
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1;

```

[3] SEQUENCE FROM N.A. Wang H., Gao X., Li L., Lou H., Huang Y., Wang B., Han J.: "Human asialoglycoprotein receptor 1 gene is expressed in SH-SY5Y human neuroblastoma cells".
PL Smith et al (2001) In the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX CARBOHYDRATE MOTIFES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE UNITS. AFTER LIGAND BINDING TO THE PROTEOP, THE RESULTING COMPLEX IS INTERNALIZED AND TRANSPORTED TO A LYSOSOME OR LYSOSOME, WHERE RECEPTOR AND LIGAND ARE DEASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE SURFACE.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARACHYMAL CELLS.
CC -1- PM: PHOSPHORYLATED ON A CYTOPLASMIC SER RESIDUE.
CC -1- MISCELLANEOUS: CALCIUM IS REPORTED FOR LIGAND BINDING.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
CC
CC EMBL: M10058; AAA51785.1;
DR EMBL: AB070943; BAB83508.1;
DR PIR: A22509; INH1.
DR HSSP: P06744; INH1.
DR Gene: BC027442; ASCB1.
DR MIM: 106460;
DR InterPro: IP001304; Lectin_C.
DR Pfam: PF00059; Lectin_C; 1.
DR SMART: S00004; CLEST; 1.
DR PROSITE: PS00425; C-TYPE_LECTIN_1.
DR PROSITE: PS00441; C-TYPE_LECTIN_2; 1.
KW Lectin; glycoprotein; receptor; endocytosis; transmembrane;
KW Calcium; signal-anchor; phosphorylation.
FT FT INIT MET 0 0
FT FT DOMAIN 1 39 CYTOPLASMIC (PROBABLE).
FT FT TRANSMEM 40 60 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT
FT FT DOMAIN 61 290 EXTRACELLULAR (PROBABLE).
FT FT DOMAIN 152 278 C-TYPE LECTIN (LONG FORM).
FT FT SITE 4 7 ENDOCYTOSIS SIGNAL (POTENTIAL).
FT FT DISULFID 153 164 BY SIMILARITY.
FT FT DISULFID 181 276 BY SIMILARITY.
FT FT CARBOHYD 254 268 BY SIMILARITY.
FT FT CARBOHYD 80 80 O-LINKED.
FT FT CARBOHYD 148 148 O-LINKED.
FT FT SEQUENCE 290 AA; 34055 MW; B1897CE30DAE1586 CP064;
Query Match 16.0%; Score 271.5; DB 1; Length 290;
Best Local Similarity 27.6%; Pred. No. 1,3e-10;
Matches 86; Conservative 45; Mismatches 112; Indels 75; Gaps 13;
QY 6 EPRVQGLGLGGLGCAVLIQIATLSPMLAGVIAIIYGVSKVPSRSSQSPQEDADAYQNI 65
DB 28 QPGLPRL-----CGPPRLALSL-----GLSLIIAYVCCYISQNSQJQHEIKGRIETP 76
QY 66 LGLAAVGLERKSLGELVYGLIIPAAVLIIPKSKQGLVYGLIIPAAVGLIIPKRS 125
DB 77 SNFLAST-----EAQVKNLSIQGVNAGKMKSL-----S 106
QY 126 KILQVAVLITPLKAAGGLDPSKIQGLVYGLIIPAAVGLIPKSKQGLVYGLIIPKRS 185
DB 107 QPGLPRL-----CGPPRLALSL-----GLSLIIAYVCCYISQNSQJQHEIKGRIETP 146
186 AFELVIRHDEPRLMTPLQCNVYKNSGQKHWKDSVIAQGVKACIVYIKCAEEDNLIQDT 245

```

DB 147 GSEKTC--CPVNWVHEHSTVWFSSTKAMADANVYPLETAHLVAVTSWEECKFYVHHI 204
QY 246 SSKNFSWNGISLDINQSTWQVDSPLSPSFQRYWNSGFPNN-----STNEIWAFFSS 269
DB 205 GGVN TWGGLHD QNGFPMKAVFSTIVETTFKRNWPEPEPLWYHGTGAGETWAFITD 259
QY 400 SGWMDNCVDNWCWK 316
DB 268 DCPMNDVQRPDPWCE 277

RESULT 11
LECL_HUMAN
ID LECL_HUMAN STANDARD: PRT: 311 AA.
AC P07307: 000448: Q03969:
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Asialoglycoprotein receptor 2 (hepatic lectin H2) (ASGP-R) (ASGPR).
EN ASGPR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
GX NCBI_TaxID=9606:
KN [1]
KP SEQUENCE FROM N.A. (ISOFORM 1).
KA MEDLINE=46016723; PubMed=1464106:
RA Spiess M., Lodish H.F.:
RT "Sequence of a second human asialoglycoprotein receptor: conservation
of two receptor genes during evolution."
RI Proc. Natl. Acad. Sci. U.S.A. 82:6465-6469(1985).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).
RC TISSUE=Liver:
KA MEDLINE=92184202; PubMed 1373982:
RA Palletta E., Stockert R.J., Racevskis J.:
RT "Differences in the abundance of variably spliced transcripts for the
second asialoglycoprotein receptor polypeptide, H2, in normal and
transfected human liver."
RI Hepatology 15:395-402(1992).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Liver:
KA Strausberg R.:
RI Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE UNITS. AFTER LIGAND
BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
DISSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
SURFACE.
CC -!- SUBUNIT: THE FUNCTIONING LIGAND-BINDING UNIT OF THIS RECEPTOR
IS THOUGHT TO BE AT LEAST A DIMER.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- ALTERNATIVE PRODUCTS: 3 isoforms: 1 (shown here), 2 and 3; are
produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CELLS.
CC -!- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL, co-sustained
by the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
or send an email to license@ebi.ac.uk).
DB EMBL: M11025; AAB59519.1;

```

```

DB EMBL: U97197; AAB58308.1;
DB EMBL: X55283; CAA38997.1;
DB EMBL: AC017251; AAH17251.1;
DB EMBL: AC017251; AAH17251.1;
DB HSR: P06734; IHL1.
DB GDB: H06734; ASGPR2.
DB MIM: 108361;
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C_1.
DR SMART: SMOU04; CLACT_1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2.
KW Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
Calcium; Signal anchor; Phosphorylation; Alternative splicing;
CYTOPLASMIC (POTENTIAL);
FT TRANSMEM 1 79
FT DOMAIN 80 311
FT DOMAIN 176 402
FT SITE 5 8
FT DISULFID 177 188
FT DISULFID 205 400
FT DISULFID 278 292
FT CARBOHYD 192 102
FT CARBOHYD 170 170
FT CARBOHYD 305 305
FT MOD_RES 12 12
FT VARSPIC 24 42
FT VARSPIC 82 86
SU SEQUENCE 311 AA; 35191 MW; 84274FECFEBA316 CRC64.

Query Match: 15.7%; Score 267; DR 1; Length 311
Best Local Similarity: 27.2%; Pct Id: 2.7%; Indels: 103
Matches 89; Conservative 49; Mismatches 103; Indels: 103

QY 6 EPRVQALALALGHALVALLSMLAVLVAIIWVSKVPSLSLEQSI
DB 47 QLAARLCSMWFS-----LALSNHIL--LTVICV-----
QY 66 TOLKAACVCHSKSKLOEYCHLOLKAACVCHPEKSKLOEYCHLTKAA
DB 77 -----TGSQSEGHGAOLQALFIRLKEAFESNF--SSHILEV--QALSTHAGS
QY 126 KQEIYQETPLKAACVCHPEKSKLOEYCHL-----ELKAAVGLH
DB 124 -----KITSIGA-----KIKQGGHAKADHALLFKLHPVLC
QY 178 QELIDKTAFLERLCHRCPKWTFQANCYENSQSRWHSVTAQOEVRAL
DB 168 H-----SNISQRTG--CPVNWVHEHSTVWFSSTKAMADANVYPLETAHLVAVTSWEECKFYVHHI
QY 246 SSKNFSWNGISLDINQSTWQVDSPLSPSFQRYWNSGFPNN-----STNEIWAFFSS
DB 221 GGVN TWGGLHD QNGFPMKAVFSTIVETTFKRNWPEPEPLWYHGTGAGETWAFITD
QY 292 EPCAFESSTG--WNHPCDEVWYK 317
DB 276 EDCVGVQFDGKWNHPCDEVWYK 302

RESULT 12
LECL_MOUSE
ID LECL_MOUSE STANDARD: PRT: 301 AA.
AC P24721:
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Asialoglycoprotein receptor 2 (hepatic lectin 2) (MHL-2)
DE (ASGPR).
EN ASGPR2 OR ASGPR-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

```

```

CX NCRL_TaxId=10090;
KN [1]
KP SOURCE: FROM N.A.
RA STRAIN=C57BL/6; TISSUE=liver;
RA MEDLINE=91027942; PubMed 2233888,
RA Sanford J P., Boyle P.;
RT "Mouse asialoglycoprotein receptor cDNA sequence (c57BL/6).
RT Receptor genes during mammalian evolution."
AL Biochim. Biophys. Acta 1087:250-261(1990)
CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC CARBOHYDRATE MOTIFIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
CC TERMINAL GALACTOSE AND N-ACETYLGLUCOSAMINE UNITS. AFTER LIGAND
CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC SURFACE.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC CELLS.
CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/submit/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X53042; CAA37211.1;
DR PIR: S13165; S13165.
DR HSP: P06734; IHL1.
DR MCD: MG1:88082; Asgr2.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C; 1.
DR SMART: SM00034; CLECT_1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.
KW Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;
KW Calcium; signal-anchor; Phosphorylation.
FT DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 59 79 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 80 301 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 169 295 C-TYPE LECTIN (LONG FORM)
FT DISULFID 170 181 BY SIMILARITY.
FT DISULFID 198 293 BY SIMILARITY.
FT DISULFID 271 285 BY SIMILARITY.
FT CARBOHYD 97 97 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 165 165 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC) (POTENTIAL).
SQ SEQUENCE 301 AA; 34907 MW; 3A29F1AFBA68F298 CRC64;
Query Match 15.28; Score 258; DB 1; Length 301;
Host Local Similarity 34.18; Pos Neg 9.8e-10;
Matches 59; Conservative 28; Mismatches 56; Indels 30; Gaps 7;

```

```

LECL_RAT
ID LECL_RAT STANDARD: PRT: 301 AA.
AC P08290;
DE 01-AUG-1988 (rel. 08, Created)
DE 01-NOV-1990 (rel. 16, last sequence update)
DE 30-MAY-2000 (rel. 39, last annotation update)
DE Asialoglycoprotein receptor L2/3 (B6.129 Lectin L2/3) (Acr. P)
DE (ASGRP).
CN ASGR2 OR ASGR 2.
OS Rattus norvegicus (rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCRL_TaxId=10116;
PN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87257885; PubMed 460647;
RA McPhaul M., Berg P.;
RT "Identification and characterization of cDNA clones encoding two
RT homologous proteins that are part of the asialoglycoprotein
RT receptor."
RL Mol. Cell. Biol. 7:1841-1847(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87250559; PubMed 3597443;
RA Halberg R.F., Wager R.E., Farrell D.C., Hildreth J., Quesenberry M.S.,
RA Loeb J.A., Holland E.C., Drickamer K.;
RT "Major and minor forms of the rat liver asialoglycoprotein receptor
RT are independent galactose-binding proteins. Primary structure and
RT glycosylation heterogeneity of minor receptor forms."
RL J. Biol. Chem. 262:9828-9836(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX SPACER-SPRINGER; TISSUE=liver;
RX MEDLINE=89170119; PubMed 3234178;
RA Santord J.P., Elliott P.W., Boyle D.;
RT "Asialoglycoprotein receptor genes are linked on chromosome 11 in the
RT mouse."
RL DNA 7:721-728(1988).
RN [4]
RP SEQUENCE OF 201-301.
RX MEDLINE=8411554; PubMed 6419386;
RA Drickamer K., Mamon J.F., Blinn G., Lemm J.O.;
RT "Primary structure of the rat liver asialoglycoprotein receptor.
RT Structural evidence for multiple polypeptide species."
RL J. Biol. Chem. 259:770-778(1984).
CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO
CC WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX
CC CARBOHYDRATE MOTIFIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES
CC TERMINAL GALACTOSE AND N-ACETYLGLUCOSAMINE UNITS. AFTER LIGAND
CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC SURFACE.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL
CC CELLS.
CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.
CC -1- IDENTIFIED: RHL-1 AND RHL-2/3, HAVING A RELATIVE ABUNDANCE OF 4:1.
CC RHL-2 AND RHL-3 ONLY DIFFERS IN THEIR CARBOHYDRATE STRUCTURES.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/submit/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: M16447; AAA42038.1;
DR EMBL: J02762; AAA41522.1;
DR EMBL: X07646; CAA30476.1;

```

RESULT 13

DR PIR: A25417; LNR12.
 DR PIR: A26888; A26888.
 DR PIR: A28462; A28462.
 DR PIR: A31601; A31601.
 DR HSSP: P06734; IHL1.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00055; Lectin_C.1.
 DR SMART: SM00034; CLECT_1.
 DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
 DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 DR Lectin: Glycoprotein; Receptor; Endocytosis; Transmembrane;
 KM Calcium; Signal-anchor; Phosphorylation.
 FT DOMAIN 1 58
 FT TRANSMEM 1 79
 FT TRANSMEM 59 79
 FT DOMAIN 80 101
 FT DOMAIN 169 295
 FT DISULFID 170 181
 FT DISULFID 194 293
 FT DISULFID 271 285
 FT CARBOHYD 97 97
 FT CARBOHYD 119 119
 FT CARBOHYD 165 165
 FT CONFLICT 153 153
 FT CONFLICT 202 202
 FT CONFLICT 260 260
 FT SHOUCE 401 AA; 34943 MM; 3C315E642D71279 CR664;
 Query Match 14 2%; Score 240; DR 1; Length 301;
 Host Local Similarity 37.4%; Pred. No. 1; 3e-08;
 Matches 49; Conservative 22; Mismatches 48; Indels 12; Gaps 5.

QY 194 GPKDPTFPQNCYFMSNSORHNHDSVTAQCEYRAQIVIKIAEONFLQIQTSKSNFMSW 253
 DB 170 GPNVWVFFGSGCYWFSRSDITVAEADQYQMEIAHLVLINSRDEOF--VYKRGAPFIH 227
 QY 254 MGLSDINQETQWGVSPSPSFQRYWNSGEPNN-----SNPFCAPFSQSG--WNNR 306
 DB 228 IGLTID--KQDSMKWVDTGYTPRSPKRN--MAFTQPDN--CGHREGSGSEDCATILSDCLMNF 284
 QY 307 GVDNFWICK 317
 DB 285 GQVNRWACER 295
 RESULT 14
 NCBI_MACMU
 ID NCBI_MACMU STANDARD: PRT: 216 AA.
 AC QYML27;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE NKG2-D type II integral membrane protein (NKG2-D activating NK
 DE receptor) (NK cell receptor D).
 GN NKG2D.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20322487; PubMed=1086118;
 RA Labonte M.L., Levy D.B., Iatvin N.L.;
 RT "Characterization of rhesus monkey CD94/NKG2 family members and
 RT identification of novel transmembrane-deleted forms of NKG2-A, B, C,
 RT and D.";
 RL Immunogenetics 51:496-499(2000).
 CC -1- FUNCTION: PLAYS A ROLE AS A RECEPTOR FOR THE RECOGNITION OF MHC
 CC CLASS I HLA-E MOLECULES BY NK CELLS AND SOME CYTOTOXIC T-CELLS.
 CC -1- SUBUNIT: CAN FORM DISULFIDE-BONDED HETERODIMER WITH CD94.
 CC -1- SUBCELLULAR LOCATION: TYPE II membrane protein.
 CC -1- TISSUE SPECIFICITY: NATURAL KILLER CELLS.

CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced by the
 CC the European Bioinformatics Institute. There are no fees
 CC use by non-profit institutions as long as its content
 CC modified and this statement is not removed. Usage by
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/seqdb/doc/seqdb.html>).
 CC
 CC EMBL: AF199943; AA074539; 1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1.
 DR SMART: SM00034; CLECT_1.
 DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 DR Receptor; Transmembrane; Multidomain family; Signal anchor;
 KM Glycoprotein; Polymorphism.
 FT DOMAIN 1 51
 FT TRANSMEM 52 72
 FT TRANSMEM 52 72
 FT DOMAIN 73 216
 FT DOMAIN 98 213
 FT DISULFID 99 110
 FT DISULFID 127 211
 FT DISULFID 184 203
 FT CARBOHYD 115 115
 FT CARBOHYD 131 131
 FT CARBOHYD 163 163
 FT CARBOHYD 202 202
 FT SHOUCE 216 AA; 25075 MM; A4483F41A00EEAC CR664;
 Query Match 13.2%; Score 224.5; DR 1; Length 216;
 Host Local Similarity 33.4%; Pred. No. 8; 7e-08;
 Matches 48; Conservative 26; Mismatches 51; Indels 5.

QY 174 QQYQELTDLKTAFERLCRCHQKIDWTFQNCYFMSNSORHNHDSVTAQCEYRAQIVIKIAEONFLQIQTSKSNFMSW 253
 DB 85 QYVYHLT --FSVQGYTRFMQYVRYRYVYTR--SKNYESAS--MS-- 227
 QY 234 TAEBONFLQIOTSRNSNFSNMSI SII NQETQWGVSPSPSFQRYWNSGEPNN-----SNPFCAPFSQSG--WNNR 306
 DB 139 SKEDDILKIVKS---YHMGVLHIIPIKSMQWFGSIIISPILLTIEMOK 284
 QY 294 CAPFSGS--GMNDRCDQVNWYIC 315
 DB 189 CALYASSFKQIEN-CSTPNIVTE 211
 RESULT 15
 LEC3_MEGRO
 ID LEC3_MEGRO STANDARD: PRT: 162 AA.
 AC P07439;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lectin BRA-3 precursor.
 OS Megabalanus rosa (Acorn barnacle).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea;
 OC Maxillopoda; Cirripedia; Thoracica; Sessilia; Balanomorpha.
 OX Balanoidae; Balanidae; Megabalanus.
 OX NCBI_TaxID=6680;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9329294; PubMed=8514190;
 RA Takematsu N., Takeda T., Kojima M., Heishi M., Muramoto K.
 RA Kamiya H., Shiba T.;
 RT "Acorn barnacle Megabalanus rosa lectin (BRA-3): cDNA clone
 RT structure and seasonal changes of mRNA and protein levels
 RT Gene 128:251-255(1993).
 RL [2]
 RP SEQUENCE OF 25-162.
 RA Muramoto K., Kamiya H.;
 RT "The amino acid sequence of a lectin of the acorn barnacle

```

RI  Mucabalanus rosa ";
KL  Biochim Biophys. Acta 874:285-295(1986).
CC  -1- FUNCTION: SUGAR-BINDING PROTEIN WHICH RECOGNIZES SPECIFIC
CC  CARBOHYDRATE STRUCTURES AND AGGLUTINATE A VARIETY OF ANIMAL
CC  CELLS BY BINDING TO CELL-SURFACE GLYCOPROTEINS AND GLYCOLIPIDS.
CC  THIS IS A CALCIUM-DEPENDENT LECTIN. INVERTEBRATE LECTINS MAY BE
CC  INVOLVED IN DEFENSE FUNCTIONS.
CC  -1- SUBUNIT: HOMOTETRAMER; DISULFIDE-LINKED.
CC  -1- TISSUE SPECIFICITY: COLELMIC FLUID.
CC  -1- MISCELLANEOUS: THIS LECTIN BINDS GALACTOSE.
CC  -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC  -----
CC  This SWISS-Prot entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcements/
CC  or send an email to license@sib-sib.ch).
CC  -----
DR  EMBL: D13302; HAA02556.1; -
DR  EMBL: D13299; HAA02556.1; JOINED.
DR  EMBL: D13300; HAA02556.1; JOINED.
DR  EMBL: D13301; HAA02556.1; JOINED.
DR  PIR: A26094; INRC3.
DR  HSSP: P23806; 1XX.
DR  InterPro: IPR001304; Lectin_C.
DR  Pfam: PF00059; Lectin_C.1.
DR  SMART: SM00034; CLECT.1.
DR  PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR  PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KW  Lectin; calcium; signal.
FT  SIGNAL      1      24
FT  CHAIN       1      162
FT  DOMAIN      25      152
FT  DISULFID    26      39
FT  DISULFID    56      150
FT  DISULFID    125     142
FT  DISULFID    157     157
FT  DISULFID    160     160
FT  VARIANT     146     146
FT  VARIANT     146     146
SO  SOURCE: 162 AA; 18328 MW; ERTFLEPQINDQBAI CPC64.

Query Match      12.9%; Score 219.5; DB 1; Length 162;
Best Local Similarity 32.5%; Pred. No. 1.3e-07;
Matches 51; Conservative 30; Mismatches 61; Indels 15; Gaps 7;

QY  172 KKGQIVYETTRDKTAPEPLCPHP-----KQWTFEGQNCYFMSNSQPNMIDSVTAQGYE- 226
DB  2 QRSFIVQAVTLVAVFAITTAECTPGNLDWQEDGHCWASTYQVRNDAQLACQTVHP 61

QY  227 -AQLVVKTAEBQNFLOLOTSRSNRPMSGMSPLNDEGTQWVDSPSPSPQRYMNSGE 285
DB  62 GAYLATIIGSOLENAFIS-ETVSNRRL-WIGLNDIDLESHYVMSNGE--ATDF-TYWSSNN 116

QY  286 PNNNSNFTAA----EFSGSAMNLNFTVFNWLCKKP 318
DB  117 PNNWENQDCGVVYDVTGQMDDDCNKNKNFLCKMP 153

```

Search completed: December 7, 2002, 11:02:24
 Job time : 30 secs

CC corresponding to the expressed proteins, and the screening of low
 CC mol.wt. medicines.

XX Sequence 296 AA;

Query Match 51.7%, Score 168, ID 23, Length 296;
 Best Local Similarity 100.0%; Pred. No. 1,2e+150;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDSKPRVQGLGILGCGHAGVALLVAVSVSPSLSGESGSDA 69

Db 1 MSDSKPRVQGLGILGCGHAGVALLVAVSVSPSLSGESGSDA 69

QY 61 IYQNTQTKAAVGLSESKSLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGL 120

Db 61 IYQNTQTKAAVGLSESKSLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGL 120

QY 121 LPEKSKLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGLP 168

Db 121 LPEKSKLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGLP 168

RESULT 4

ID AAG79093 standard; Protein: 296 AA

XX AAG79093;

XX 10-DEC-2001 (first entry)

DE Amino acid sequence of a splice variant of human DC-SIGN.

XX Human: receptor; DC-SIGN: dendritic cell; T lymphocyte; HIV;

XX gp120; C-type lectin; ICAM3; HIV entry; T cell; macrophage;

XX HIV infection; splice variant.

XX Homo sapiens.

XX W0200164752-A2.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001W0-US06322.

XX 02-MAR-2000; 2000US-0517605.

XX (UNYNY) UNIV NEW YORK STATE.

XX (UNYNY) UNIV NJUMCGEN.

XX Liltman DK, Kwon D, Van Kooyk Y, Neijtenbeek T;

XX WPI: 2001-602565/68.

XX N-PSDB: AAI65469.

XX Example 1: Page 130-131; 131pp; English.

XX The present sequence represents a splice variant of a human receptor

XX designated DC-SIGN. DC-SIGN is specifically expressed on dendritic cells

XX and facilitates infection of T lymphocytes with HIV. DC-SIGN is identical

XX to a HIV-1 gp120-binding C-type lectin. DC-SIGN binds ICAM-3 (which is

XX expressed constitutively on T lymphocytes) with high affinity. The

XX specification describes an antibody which is specific for an antigenic

XX fragment of gp120. This antigenic fragment binds to DC-SIGN or is exposed

XX upon gp120 binding of DC-SIGN due to conformational change.

XX The antibody inhibits the trans enhancement of HIV entry into a T

XX cell or macrophage facilitated by dendritic cells. The antibody is

XX useful to treat or prevent HIV infection.

Query Match 51.7%, Score 168, ID 23, Length 296;
 Best Local Similarity 100.0%; Pred. No. 1,2e+150;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSDSKPRVQGLGILGCGHAGVALLVAVSVSPSLSGESGSDA 69

Db 1 MSDSKPRVQGLGILGCGHAGVALLVAVSVSPSLSGESGSDA 69

QY 61 IYQNTQTKAAVGLSESKSLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGL 120

Db 61 IYQNTQTKAAVGLSESKSLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGL 120

QY 121 LPEKSKLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGLP 168

Db 121 LPEKSKLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGLP 168

RESULT 4

ID ABB09715 standard; Protein: 296 AA.

XX ABB09715;

XX 11-JUN-2002 (first entry)

DE Amino acid sequence of human polypeptide HP01347.

XX Antibody; antigen; transmembrane domain protein; HP01347

XX Homo sapiens.

XX W0200208416-A1.

XX 31-JAN-2002.

XX 24-JUL-2001; 2001W0-JP06371.

XX 24-JUL-2000; 2000JP-0222743.

XX 24-AUG-2000; 2000JP-0254407.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Kato S, Nagata N, Fujimura N, Kobayashi M, Ito K

XX WPI: 2002-195877/25.

XX N-PSDB: ABL41993.

XX Antibody preparation by inoculation of an animal with a

XX transmembrane domain for use as drugs, diagnostic reagents

XX laboratory reagents.

XX Example; Page 23-27; 45pp; Japanese.

XX The specification describes a method of antibody preparation

XX method comprises inoculating an animal with a vector carrying

XX a fusion protein having an antigen protein fused to the

XX side (extracellular) of a transmembrane domain protein

XX N-terminal side of which is intracellular, and then is

XX purifying the antibody from the animal. The antibodies

XX as drugs, vaccines, diagnostic reagents and laboratory

XX used in the course of the invention.

XX Sequence 296 AA;

XX Query Match 51.7%, Score 168, ID 23, Length 296;

XX Best Local Similarity 100.0%; Pred. No. 1,2e+150;

XX Matches 168; Conservative 0; Mismatches 0; Indels 0;

XX QY 1 MSDSKPRVQGLGILGCGHAGVALLVAVSVSPSLSGESGSDA 69

XX Db 1 MSDSKPRVQGLGILGCGHAGVALLVAVSVSPSLSGESGSDA 69

XX QY 61 IYQNTQTKAAVGLSESKSLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGL 120

XX Db 61 IYQNTQTKAAVGLSESKSLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGL 120

XX QY 121 LPEKSKLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGLP 168

XX Db 121 LPEKSKLQEIYQELTQKAAVGLPEPSKSLQEIYQELTQKAAVGLP 168

Db 1 MSDSKPRVQGLIAGTIGHALVQLSPMLAGVLAIVGVSPSSI SQSEQDA 60
 QY 61 IYQMLTQKAAVGLSKSLQGLTYQELTQKAAVGLPEKSKLQETLYELTRKAAGV 120
 Db 61 IYQMLTQKAAVGLSKSLQGLTYQELTQKAAVGLPEKSKLQETLYELTRKAAGV 120
 QY 121 LPEKSKLQETLYQELTRKAAVGLPEKSKLQETLYQELTRKAAVGLP 168
 Db 121 LPEKSKLQETLYQELTRKAAVGLPEKSKLQETLYQELTRKAAVGLP 168
 RESULT 5
 AAB94720
 ID AAB94720 standard; Protein: 216 AA.
 AC AAB94720;
 XX 26-JUN-2001 (first entry)
 DE Human protein sequence SEQ ID NO:15736.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 FN EP1074617-A2.
 XX 07-FEB-2001.
 XX 28-JUL-2000; 2000EP-0116126.
 XX 29-JUL-1999; 99JP-0248036.
 XX 27-AUG-1999; 99JP-0300253.
 XX 11-JAN-2000; 2000JP-0118774.
 XX 02-MAY-2000; 2000JP-0183767.
 XX 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 FA Ota I, Isogai T, Nishiyama T, Hayashi K, Saito K, Yamamoto J;
 FI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 FI full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 XX full-length cDNAs -
 PS Claim 8; SEQ ID 15736; 2537pp + CD ROM; English.
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAB94720 to AAB19428 and
 CC AAB19433 to AAB18742 represent human cDNA sequences; AAB92446 to
 CC AAB95894 represent human amino acid sequences; and AAB13529 to AAB19432
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 216 AA:
 SQ Query Match 29.8%; Score 97; DB 22; Length 216;
 CC Best local similarity 100.0%; Pred. No. 1,46-83;
 CC Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 72 VQPISEKSKLQETLYQELTQKAAVGLPEKSKLQETLYQELTRKAAVGLPEKSKLQETLY 131
 Db 44 VQPISEKSKLQETLYQELTQKAAVGLPEKSKLQETLYQELTRKAAVGLPEKSKLQETLY 103
 QY 132 QELTRKAAVGLPEKSKLQETLYQELTRKAAVGLP 168
 Db 104 QELTRKAAVGLPEKSKLQETLYQELTRKAAVGLP 140
 RESULT 6
 AAR32188
 ID AAR32188 standard; Protein: 404 AA.
 AC AAR32188;
 XX 08-JUN-1993 (first entry)
 DE Sequence of a non CD4 glycoprotein gp120 receptor protein.
 XX Surface cell protein, glycoprotein receptor protein, gp120 receptor,
 XX HIV; diagnosis; therapy.
 XX Homo sapiens.
 FN W09301820-A.
 XX W09301820-A.
 XX 04-FEB-1993.
 XX 16-JUL-1992; 92WO-US05985.
 XX 16-JUL-1991; 91US-0731214.
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 FA Curtis BM;
 FI WPI: 1993-058515/07.
 XX N-PSDB; AA036648.
 XX Inhibition of non-CD4 mediated HIV infection dependent on novel
 PT gp-120 receptor, for diagnosing and treating HIV infection
 XX Disclosure; Figure 3A; 44pp; English.
 XX The gp120 receptor protein is present on placental, muscle, neural,
 CC brain, dendritic or mucosal cells. It may be produced recombinantly
 CC by cloning cDNA isolated from a library of recombinant placental
 CC genes. Binding of gp120 to its receptor is inhibited by specific
 CC carbohydrate, plant lectins and specific antibodies. It corresp.
 CC to a non-CD4 mammalian cell surface protein, contg. ca. 400 AAs and
 CC having a mol. wt. of ca. 15 kD, and a binding affinity for gp-120 of
 CC kd = 1.32-2 nM.
 XX Sequence 404 AA:
 SQ Query Match 24.0%; Score 76; DB 14; Length 404;
 CC Best local similarity 100.0%; Pred. No. 2,56-65;
 CC Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 72 QPVALYQMLTQKAAVGLSKSLQGLTYQELTQKAAVGLPEKSKLQETLYQELTRKAA 117
 Db 74 QPVALYQMLTQKAAVGLSKSLQGLTYQELTQKAAVGLPEKSKLQETLYQELTRKAA 133
 QY 118 VQPISEKSKLQETLYQELT 135
 Db 134 VQPISEKSKLQETLYQELT 151

db 134 VQELPEKSKLOEIYQELT 151

RESU1.9
AAG79086 standard: Protein: 404 AA.
XX
XX AAG79086;
XX
XX 10-DEC-2001 (first entry)
XX
XX Human DC-SIGN, a dendritic cell-specific C-type lectin.
DE
XX
XX Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV;
KM qpi20; C type lectin; ICAM3; HIV entry; T cell; macrophage;
KM HIV infection.
XX
XX Homo sapiens.
OS
XX W0200164752-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-0536322
XX
XX 02-MAR-2000; 2000US-0517605
XX
XX (UYNY) UNIV NEW YORK STATE.
PA (UYNY-) UNIV NITMAGEN
XX
XX Littman DP, Kwau D, Van Kowyk Y, Gelfenbeck T;
PI WPI: 2001-602565/68
XX
XX N-PSDB: AA165461.
DR
XX An antibody for the treatment or prevention of HIV-infection comprises
PT a qpi20 portion which binds to DC-SIGN or is exposed upon qpi20 binding
PT of DC-SIGN due to concomitant conformational change.
XX
XX Example 1: Page 114-115, 131pp; English.
PS
XX The present sequence represents a human receptor designated DC-SIGN, that
CC is specifically expressed on dendritic cells and facilitates infection
CC of T lymphocytes with HIV. DC-SIGN is identical to a HIV-1
CC qpi20 binding C-type lectin. DC-SIGN binds ICAM 3 (which is expressed
CC constitutively on T lymphocytes) with high affinity. The specification
CC describes an antibody which is specific for an antigenic fragment of
CC qpi20. This antigenic fragment binds to DC-SIGN or is exposed upon
CC qpi20 binding of DC-SIGN due to concomitant conformational change.
CC The antibody inhibits the trans enhancement of HIV entry into a T
CC cell or macrophage, facilitated by dendritic cells. The antibody is
CC useful to treat or prevent HIV infection
XX
XX Sequence 404 AA:
SQ

Query Match: 24.0%; Score 78; DB 23; Length 404;
Best Local Similarity 100.0%; Pred. No. 2, 5e-65;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

db 74 QQAIVYQNLTKAAGVSEKSKLOEIYQELTQKAAGVSEKSKLOEIYQELTRKAA 133

QY 58 QQAIVYQNLTKAAGVSEKSKLOEIYQELTQKAAGVSEKSKLOEIYQELTRKAA 117
db 118 VQELPEKSKLOEIYQELT 135
db 134 VQELPEKSKLOEIYQELT 151

RESULT 10
AA08025
ID AA08025 standard: Protein: 404 AA.
XX
XX AA08025;
AC AA08025;

XX 05-JUN-2002 (first entry)
XX
XX Mouse C1S1-B7 ORF protein.
DE
XX
XX Nucleic acid library: immune response; asthma; COPD;
KM airway hyperresponsiveness; bronchial; allergic asthma;
KM signature sequence; SS; chronic obstructive pulmonary disease;
KM allergic disease; rhinitis; atopic dermatitis; urticaria;
KM autoimmune disease; multiple sclerosis; inflammatory bowel disease;
KM allograft rejection; infectious disease;
KM calcium activated chloride channel.
XX
XX Mus sp.
OS
XX W020021466-A2.
XX
XX 21-FEB-2002.
XX
XX 16-AGC-2001; 2001WO-NL00610.
XX
XX 16-AGC-2000; 2000EP-0202867.
XX
XX (UYNY) UNIV SUNNY DRECHT.
PA
XX
XX Groot PC, Van Bergenhenegouwen BJ, Van Oosterhout AJM;
PI WPI: 2002-241886/29.
XX
XX Nucleic acid library comprising genes which are capable of initiation,
PT progression and suppression of an immune response, especially an immune-
PT response observed with airway hyper-responsiveness of asthma.
XX
XX Example 10; Fig 11; 120pp; English.
PS
XX The invention relates to a nucleic acid library comprising genes or
CC their fragments which are capable of modulating an immune response
CC observed with airway hyper-responsiveness and/or bronchial/vascular
CC manifestations of asthma. Also included are a method for modulating an
CC immune response of an individual comprising modulating a gene comprising
CC a nucleic acid at least functionally equivalent to a nucleic acid
CC identifiable by a signature sequence (SS) given in the specification such
CC as p1-S1-A11, S101-A10, S102-111, S101-A12, and p1-S0-R1-B7, a
CC substance (for use as a modulator) capable of modulating a gene
CC comprising a nucleic acid at least functionally equivalent to a nucleic
CC acid identifiable by SS and the use of a prophylactic substance derived
CC from a nucleic acid at least functionally equivalent to a nucleic acid
CC identifiable by SS for the production of an antagonist (for use as a
CC medicament) against the substance. The antagonist and substance are
CC useful for the treatment of an immune response observed with airway
CC hyper-responsiveness and/or bronchial/vascular manifestations of asthma
CC The method is useful for modulating the above immune response, where the
CC gene encodes a gene product capable of modulating the immune response
CC The substance is useful for treating an immune response, particularly
CC asthma, chronic obstructive pulmonary disease (COPD), allergic diseases
CC (rhinitis, atopic dermatitis, urticaria), autoimmune diseases (e.g.
CC multiple sclerosis), inflammatory bowel disease, allograft rejection and
CC infectious disease. The present sequence is a mouse or human
CC protein encoded by a signature sequence gene or its homologue/functionally
CC equivalent.
XX
XX Sequence 404 AA:
SQ

Query Match: 24.0%; Score 78; DB 23; Length 404;
Best Local Similarity 100.0%; Pred. No. 2, 5e-65;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

db 74 QQAIVYQNLTKAAGVSEKSKLOEIYQELTQKAAGVSEKSKLOEIYQELTRKAA 133

QY 58 QQAIVYQNLTKAAGVSEKSKLOEIYQELTQKAAGVSEKSKLOEIYQELTRKAA 117
db 118 VQELPEKSKLOEIYQELT 135
db 119 VQELPEKSKLOEIYQELT 151


```

RESULT 14
AA025647
ID ABR25647 standard; Protein: 50 AA.
XX
AC ABR25647;
XX
ID 23-JAN-2002 (first entry)
XX
DE Protein #7646 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN M0200157274-A2.
XX
ID 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-0500647.
XX
PR 04-PEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0638408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15, SEQ ID NO 27417, 530pp, English
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA11533:ABA11035). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and identifying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published\_pat\_sequences.
XX
SQ Sequence 50 AA:
XX
Query Match 15.4%; Score 50; DB 22; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.2e-19;
Matches 50; Conservative 0; Mismatches 0; Indels 6; Gaps 0;
XX
CY 189 ETCRHTETKWTFFGNCYFMNSGPNMHISVATCGFVPAQIVIKTAFRQ 238
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
ID 1 RUCRHTETKWTFFGNCYFMNSGPNMHISVATCGFVPAQIVIKTAFRQ 50
XX
RESULT 14
AA062914
ID AAM62914 standard; Protein: 50 AA.
XX
AC AAM62914;
XX
ID 05-NOV-2001 (first entry)
XX

```

```

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35019.
XX
XX Human brain expressed exon gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN M0200157275-A2.
XX
ID 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-0500647.
XX
PR 04-PEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0638408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 35019; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
SQ Sequence 50 AA:
XX
Query Match 15.4%; Score 50; DB 22; Length 50;
Best Local Similarity 100.0%; Pred. N 1.2e-19;
Matches 50; Conservative 0; Mismatches 0; Indels 6; Gaps 0;
XX
CY 189 ETCRHTETKWTFFGNCYFMNSGPNMHISVATCGFVPAQIVIKTAFRQ 238
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
ID 1 ETCRHTETKWTFFGNCYFMNSGPNMHISVATCGFVPAQIVIKTAFRQ 50
XX
RESULT 15
AA075726
ID AAM75726 standard; Protein: 50 AA.
XX
AC AAM75726;
XX
ID 06-NOV-2001 (first entry)
XX
DE Human brain expressed probe encoded protein SEQ ID NO: 36032.
XX
XX Human brain expressed exon gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN M0200157276-A2.
XX
ID 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-0500668.
XX

```

04-FEB-2000: 2000US-0180312
 PR 26-MAY-2000: 2000US-0207456.
 PR 10-JUN-2000: 2000US-0608408.
 PR 03-AUG-2000: 2000US-0632466.
 PR 21-SEP-2000: 2000US-0234687.
 PR 27-SEP-2000: 2000US-0246459.
 PR 04-OCT-2000: 2000US-0024263.
 (MOLF-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI: 2001-488900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 XX Example 4: SEQ ID NO: 3632, 658bp + Sequence Listing; English.
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 SO Sequence 50 AA:
 Query Match 15.4%; Score 50; DR 22; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1,2e-39;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 YQ 189 RLCRHCPRKMTFFGNGCYFMSNSGRNMHSVIAAGCVRAQIVIKIAEEQ 238
 Db 1 RLCRHCPRKMTFFGNGCYFMSNSGRNMHSVIAAGCVRAQIVIKIAEEQ 50
 AAM35835
 ID AAM35835 standard; Protein: 50 AA.
 AC AAM35835;
 XX 17-OCT-2001 (first entry)
 DE Peptide #9872 encoded by probe for measuring placental gene expression.
 XX Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 OS Homo sapiens.
 XX
 PN W0200157272-A2.
 PD 09-AUG-2001.
 XX
 XX 30-JAN-2001: 2001WO-US00663.
 PR 04-FEB-2000: 2000US-0180312.
 PR 26-MAY-2000: 2000US-0207456.
 PR 10-JUN-2000: 2000US-0608408.
 PR 03-AUG-2000: 2000US-0632466.
 PR 21-SEP-2000: 2000US-0234687.
 PR 27-SEP-2000: 2000US-0236359.
 PR 04-OCT-2000: 2000US-0024263.
 (MOLF-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI: 2001-488900/53.
 XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -
 XX
 XX Claim 27: SEQ ID No 36104: 654bp; English.
 PS
 XX The present invention relates to single exon nucleic acid
 CC see A13315-A157546). The present sequence is a peptide
 CC such probe. The probes are useful for producing a microarray
 CC predicting, measuring and displaying gene expression in
 CC from human placenta. The probes are useful for antenatal
 CC human genetic disorders.
 XX
 SO Sequence 50 AA:
 Query Match 15.4%; Score 50; DR 22; Length 50;
 Best Local Similarity 100.0%; Pred. No. 1,2e-39;
 Matches 50; Conservative 0; Mismatches 0; Indels 0;
 YQ 189 RLCRHCPRKMTFFGNGCYFMSNSGRNMHSVIAAGCVRAQIVIKIAEEQ
 Db 1 RLCRHCPRKMTFFGNGCYFMSNSGRNMHSVIAAGCVRAQIVIKIAEEQ
 AAG45211
 ID AAG45211 standard; Peptide: 50 AA.
 AC AAG45211;
 XX 19-AUG-2002 (first entry)
 DE Human peptide encoded by genome-derived single exon probe
 XX
 XX Human; single exon probe; asthma; lung cancer; COPD; chronic
 KW obstructive pulmonary disease; interstitial lung
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hereditary spherocytosis; lymphangioleiomyomatosis; Karyo-
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary disease;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 KW
 OS Homo sapiens.
 XX
 PN W0200186003-A2.
 PD 15-NOV-2001.
 XX
 XX 30-JAN-2001: 2001WO-US00665.
 PR 04-FEB-2000: 2000US-180312P.
 PR 26-MAY-2000: 2000US-207456P.
 PR 10-JUN-2000: 2000US-0608408.
 PR 03-AUG-2000: 2000US-0632466.
 PR 21-SEP-2000: 2000US-244647P.
 PR 27-SEP-2000: 2000US-246359P.
 PR 04-OCT-2000: 2000US-0024263.
 (MOLF-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI: 2002-114183/15.
 XX Specifically addressable set of single exon nucleic acid probes
 PT measure gene expression in human lung samples -
 XX
 XX Claim 27: SEQ ID No 34876: 644bp; English.
 PS
 XX The invention relates to a spatially-addressable set of
 CC nucleic acid probes for measuring gene expression in a
 CC from human lung comprising single exon nucleic acid probes
 CC 12614 nucleic acid sequences mentioned in the specification

complements of the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote, and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above, and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarray having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hutchinson-Podak syndrome, sarcoidosis, Pseudomonas haemolyticus, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagenor syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention.

Note: The sequence data for this patent did not form part of the prior art specification, but was obtained in electronic format directly from Wipo at http://wipo.int/pub/published_pcl_sequences.

Sequence 50 AA:
 Query Match 15.4%; Score 50; DR 23; Length 50;
 Best Local Similarity 100.0%; Pred. No 1 2e-19;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 189 RLCRHCPKMTFFGNCYFMSNSGRNMHSYACGEVAVLVKTADEQ 238
 1 RLCRHCPKMTFFGNCYFMSNSGRNMHSYACGEVAVLVKTADEQ 50

RESULT 18
 AAY72948
 ID AAY72948 standard; Protein: 238 AA.

XX AAY72948;

DT 13-JUN-2001 (first entry)

DE Mouse type II integral membrane protein, CIRE.

XX Mouse; type II integral membrane protein; CIRE; anti-influenza vaccine; immunosuppressive; autoimmune disorder; allergy; viral infection; APC; antigen presenting cell; bacterial infection; macrophage; dendritic cell; etc.

XX Mus musculus.

PN W020011986-9-A1.

PD 22-MAR-2001.

PF 13-SEP-2000; 2000WO-AU01096.

PR 13-SEP-1995; 99AU-0002788.

PA (COEN) CORNELL GREENSLAND INST MEDICAL RES
 XX Caminschi I, Wright MD, Shortman KD;
 PI WPI: 2001-244786/25.
 DR N-PSDB: AAD04221.

XX Novel type II integral membrane protein, CIRE, preferentially expressed in dendritic cells, macrophages and their precursors stimulatory to T cells; useful for identifying compounds having immunoregulatory activity.

Claim 1: Fig 1; 53pp; English.

The present sequence is mouse type II integral membrane protein (designated as CIRE) which is preferentially expressed in dendritic cells, macrophages and their precursors stimulatory to T cells. CIRE protein has a N-terminal cytoplasmic domain, transmembrane domain and a C-type lectin domain in the extracellular region. CIRE is useful for screening immunoregulatory compounds and as a marker for dendritic cell (DC) population. CIRE specific ligand such as monoclonal antibodies are useful for isolating an antigen presenting cell (APC) from a biological sample. CIRE DNA is also used as tools to analyse the properties and functions of the CIRE gene/protein. For example, the DNA molecules are used to generate animal models, which lack functional CIRE genes and to isolate regulatory regions of the CIRE gene. Such regulatory regions are used to selectively express exogenous genes in DC or APC. Ligands are useful for modulating immune responses by interacting with the function, migration or maturation of DC or APC. The ligands bind to and interfere with the function of myeloid dendritic cells such that antigen processing is modulation by lymphoid DC which leads to immune suppression and anergy for treating allergies and autoimmune disorders. The ligands are also useful to block or enhance the interaction of viruses or bacteria with DC, for improving resistance to infection. These ligands are also used to target molecules such as vaccine components, to DC or APC.

Sequence 238 AA:

Query Match 2.8%; Score 9; DR 22; Length 238;
 Best Local Similarity 100.0%; Pred. No 3,1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 85 YQHTDQKA 93
 91 YQHTDQKA 99

RESULT 19
 AAM25722
 ID AAM25722 standard; Protein: 394 AA.

XX AAM25722;

DT 08-MAY-1998 (first entry)

DE Human partial beta meltrin protein fragment 2.

XX Meltrin; myotube formation; muscle tissue; therapy; membrane protein; myoblast; adhesion; fusion; medicinal; muscle disorder; human.

XX Homo sapiens.

PN W09731109-A1.

PD 28-AUG-1997.

PF 17-OCT-1996; 96WO-0004017.

PR 23-FEB-1996; 96JP-0061756.

PA (MORI) MOCHIDA PHARM CO LTD.
 XX Fujisawa A, Mizushima C, Ogawa N, Shirakawa K, Yamakawa T;

XX WP: 1997-435161/40.
 DR N-PSDB: AAT6014
 XX Meltrin proteins involved with myotube formation in muscle - contain
 PT metalloproteinase and disintegrin domains
 XX
 PS Claim 5; Fig 17; 262pp; Japanese.
 CC AAM25721 and AAM25722 represent fragments of a new human membrane
 CC protein, beta meltrin, which participates in the adhesion and fusion of
 CC myoblasts in the course of myotube formation in muscle tissue. Meltrins
 CC and their partial peptides can be used in the study of myotube formation,
 CC and in medicinal compositions for muscle disorders. Meltrins contain an
 CC intracellular domain, a membrane-penetrating domain, a fusion peptide
 CC domain, a disintegrin domain, a metalloproteinase domain, a precursor
 CC domain and a cysteine rich domain although these are not shown in the
 CC specification. Meltrin antagonists can be used as bone absorption
 CC inhibitors, e.g. for the treatment of osteoporosis and hypercalcaemia
 CC and for prevention of cancer cell infiltration.
 CC
 XX Sequence 394 AA:
 SU
 Query Match 2.88; Score 9; DB 18; Length 394;
 Best Local Similarity 100.0%; Pred. No. 5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 YY 34 AGVLVAILV 42
 DB 333 AGVLVAILV 341
 RESULT 20
 AAM01825
 ID AAM01825 standard; Protein; 751 AA.
 XX
 AC AAM01825;
 XX
 DT 27-MAY-1997 (first entry)
 XX
 DE Human metalloproteinase.
 XX
 KW Metalloproteinase; infection; inflammation; allergy; therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 26..240
 FT /label= Catalytic_domain
 XX
 PN W09709420-A2.
 XX
 PD 13-MAR-1997.
 XX
 PF 05-SEP-1996; 96MO-GH02180.
 XX
 PR 11-JUN-1996; 96GB-0012145.
 PR 05-SEP-1995; 95GB-0018026.
 PR 20-OCT-1995; 95GB-0021496.
 PR 20-OCT-1995; 95GB-0021497.
 PR 21-DEC-1995; 95GB-0026230.
 XX
 PA (GILT) CELLTech THERAPEUTICS LTD.
 XX
 PI Docberry AIF, Slocombe PM;
 XX
 DR WP: 1997-192896/17.
 DR N-PSDB: AAT62559.
 XX
 PT DNA encoding human metalloproteinase used to develop prods. for
 PT treating, e.g. infectious diseases, inflammatory disease or allergy
 XX
 PS Claim 5; page 5; 17pp; English.

XX
 CC A novel human metalloproteinase (AAM01825), or its catalytic
 CC domain, can be expressed in prokaryotic or eukaryotic host cells
 CC an isolated gene sequence (see also AAT62558-59). The
 CC metalloproteinase can be used to screen for cpds. that
 CC are active, and to raise antibodies. Such cpds. and
 CC are useful for the in vivo treatment of diseases or dis-
 CC related to over or under regulated prodn. of the enzyme
 XX
 SU Sequence 751 AA:
 YY 34 AGVLVAILV 42
 DB 530 AGVLVAILV 538
 RESULT 21
 AA014366
 ID AA014366 standard; Protein; 751 AA.
 XX
 AC AA014366;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human metalloproteinase-like/disintegrin-like protein #1.
 XX
 KW Human; metalloproteinase-like/disintegrin-like protein; angio-
 KW genesis; endothelial cell migration; integrin-like
 KW disorder; cancer; inflammation; osteoporosis; osteo-
 KW cytolysis; antiinflammatory; osteopathy; vasculopathy;
 KW vulnerability; antiinflammatory; osteopathy; vasculopathy; pro-
 KW antiinflammatory; analgesic; thrombosis; tissue repair; neuro-
 KW infection.
 XX
 OS Homo sapiens.
 XX
 PN W0200210406-A2.
 XX
 FD 07-FEB-2002.
 XX
 PF 27-JUL-2001; 2001WO-US23784.
 XX
 FP 28-JUL-2000; 2000US 221848P.
 XX
 PA (IMMUNEX CORP.
 XX
 PI Dubose RF, Wiley SP, Black KA;
 XX
 DR WP: 2002-195965/25.
 XX
 PT Novel purified human metalloproteinase-disintegrin poly-
 PT for treating ocular and inflammatory disorders, osteopor-
 PT stenosis, thrombosis, and chronic pain conditions
 XX
 FS Claim 1, Fig 1; 101pp; English.
 XX
 CC The present invention provides a number of human metallo-
 CC /disintegrin-like (MPD) proteins and their fragments. These
 CC are used to identify treatments for ocular disorders, malignant
 CC conditions, inflammatory diseases, osteoporosis and other
 CC mediated by accelerated bone resorption, restenosis, throm-
 CC platelet activation, recruitment or aggregation, thrombo-
 CC requiring tissue repair or wound healing, endothelial mal-
 CC angiogenesis, allergies, reproductive, neurological and
 CC conditions, infectious, cardiovascular disorders such as
 CC intarction, chronic pain conditions, endocrine system dis-
 CC gastrointestinal system disorders, genitourinary system
 CC anaemia and haematological disorders. The present sequen-
 CC of the invention.

XX Sequence 751 AA:
 SO Query Match 2.8%; Score 9; DB 23; Length 751;
 Best Local Similarity 100.0%; Pred No. 9.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AGVIVALLV 42
 |||||
 DB 530 AGVIVALLV 538

RESULT 22
 AAB50942
 ID AAB50942 standard: peptide; 778 AA.
 XX AAB50942:
 AC AAB50942:
 XX 20-MAR-2001 (first entry)
 DT
 XX ADAM gene #1 peptide #1
 DE
 XX ADAM, disintegrin, metalloproteinase, food additive, breast cancer;
 KM ovarian.
 KW
 XX Homo sapiens.
 OS
 XX MO200073323-A2.
 PN
 XX 07-DEC-2000.
 PD
 XX 25-MAY-2000; 2000MO-US1410A
 PE
 XX 27-MAY-1999; 990S-0136388.
 PR 69-JUL-1999; 99US-0142949.
 PP 28-JAN-2000; 2000US-0178717.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Wei P, Ni J, Hastings GA, Shi Y;
 PI WPI; 2001-015507/02.
 DR
 XX Seven nucleic acid molecules encoding ADAM polypeptides containing a
 PT disintegrin and metalloprotease domain, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases -
 XX
 PS Disclosure, Page 12, 287pp, English.
 CC
 XX The present invention relates to seven members of the ADAMs (proteins
 CC which contain A Disintegrin And Metalloprotease domain) protein family.
 CC The ADAMs contain A Disintegrin And Metalloprotease domain, as a food
 CC additive or preservative, for chromosome identification, as probes
 CC for diagnosing a disorder related to the female reproductive system,
 CC particularly breast and/or ovary cancer. They are also useful in the
 CC gene therapy of breast and ovarian cancer.
 CC
 XX
 XX Sequence 778 AA:
 SO Query Match 2.8%; Score 9; DB 22; Length 778;
 Best Local Similarity 100.0%; Pred No. 9.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AGVIVALLV 42
 |||||
 DB 705 AGVIVALLV 713

RESULT 23
 AA014381
 ID AA014381 standard: Protein; 811 AA
 XX

AC AA014381:
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human metalloproteinase-11;disintegrin like protein #25.
 XX
 KM Human; metalloproteinase like/disintegrin like protein; MPD;
 KM angiogenesis, endothelial cell migration, integrin ligand binding,
 KM ocular disorder; cancer; inflammation; osteoporosis; restenosis;
 KM cysticercus; antiinflammatory; osteopathic; vasculopathy; thrombolytic;
 KM valine; antiatherogenic; antibacterial; vitreous; proteoglycan; cardiac;
 KM antianemic; analgesic; thrombosis; tissue repair; neural disorder;
 KM infection.
 KW
 XX
 XX Homo sapiens.
 OS
 XX MO200210406 A2.
 FN
 XX 07-FEB-2002.
 TD
 XX 27-JUL-2001; 2001MO-US23724.
 PF
 XX 28-JUL-2000; 2000US-221848P.
 PE
 XX (IMV) IMMUNEX CORP.
 PA
 XX Dubose PE, Wiley SR, Black RA;
 PI WPI; 2002-195965/25.
 DR
 XX Novel purified human metalloproteinase-disintegrin polypeptide, useful
 PT for treating cancer and inflammatory disorders, osteoporosis, cancer,
 PT restenosis, thrombosis, and chronic pain conditions -
 PT
 XX
 PS Claim 1, Fig 1, 101pp; English.
 CC
 XX The present invention provides a number of human metalloproteinase-like
 CC /disintegrin like (MPD) proteins and their fragments. These can be used
 CC to identify treatments for ocular disorders, malignant and metastatic
 CC conditions, inflammatory diseases, osteoporosis and other conditions
 CC mediated by accelerated bone resorption, restenosis, inappropriate
 CC platelet activation, recruitment or aggregation, thrombosis, conditions
 CC requiring tissue repair or wound healing, endothelial migration,
 CC angiogenesis, allergies, reproductive, neurological and vascular
 CC conditions, infections, cardiovascular disorders such as myocardial
 CC infarction, chronic pain conditions, endocrine system disorders,
 CC gastrointestinal system disorders, genitourinary system disorders,
 CC anemia and hematological disorders. The present sequence is a protein
 CC of the invention.
 CC
 XX
 XX Sequence 811 AA:
 SO Query Match 2.8%; Score 9; DB 23; Length 811;
 Best Local Similarity 100.0%; Pred No. 9.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AGVIVALLV 42
 |||||
 DB 629 AGVIVALLV 637

RESULT 24
 AA074756
 ID AA074756 standard: Protein; 879 AA.
 XX AA074756:
 AC AA074756:
 XX 09-APR-2002 (first entry)
 DT
 XX Human protease PRPS-16 protein sequence.
 DE
 XX Human; protease; PRPS; gastrointestinal, Crohn's disease, cancer,
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
 KW

inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 cell proliferative disorder; developmental disorder; epilepsy;
 Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
 reproductive disorder; endometriosis.
 OS Homo sapiens.
 FN W0200198468-A2.
 XX
 XX 27-DEC-2001.
 XX
 PF 13-JUN-2001; 2001WO-0819178
 XX
 XX 16-JUN-2000; 2000US-212336P.
 PR 22-JUN-2000; 2000US-213956P.
 PR 29-JUN-2000; 2000US-215396P.
 PR 07-JUL-2000; 2000US-216821P.
 PR 14-JUL-2000; 2000US-218946P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-Young J, Tribouley CM;
 Deleage AM, Haughn MK, Nguyen DB, Lee EA, Hatalia A, Khan FA;
 Wadia NK, Yao MG, Lu DM, Patterson C, Tang YT, Walsh RT;
 PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;
 PI Kalliek DA;
 XX
 DR WPI: 2002-090437/12.
 DP N-PSNR: AHK12899.
 XX
 PT Twenty one human proteases (referred to as PTRS-1 to PTRS-21), useful
 PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell
 PT proliferative (e.g. cancer) disorders -
 PS
 PS Claim 1; Page 151-153; 177pp; English.
 XX
 CC The present invention relates to twenty one new human proteases,
 CC referred to as PTRS-1 to PTRS-21. The PTRS polynucleotides and
 CC polypeptides of the invention are useful in the diagnosis, treatment and
 CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and
 CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and
 CC myocardial infarction, autoimmune/inflammatory e.g. acquired
 CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell
 CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker
 CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.
 CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and
 CC endometriosis disorders. Numerous other examples of each disorder are
 CC given in the specification. The present protein sequence represents
 CC the human protease PTRS-16 protein of the invention.
 XX
 SQ Sequence 879 AA;
 Query Match 2.8%; Score 9; DB 23; Length 879;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 34 AGVLVAIIIV 42
 IIIIIIII
 DB 629 AGVLVAIIIV 637
 RESULT 25
 AAB50935 standard; protein; 918 AA.
 XX
 AC AAB50935;
 XX
 DT 20-MAR-2001 (first entry)
 XX
 DE ADAM protein #1.
 XX
 XX ADAM; disintegrin; metalloprotease, food additive; breast cancer;

XX ovarian.
 XX
 OS Homo sapiens.
 XX
 PN W0200073323-A2
 XX
 PD 07-DEC-2000.
 XX
 PF 25-MAY-2000; 2000WO-US1408.
 XX
 PR 27-MAY-1999; 99US-013648P.
 PR 09-JUL-1999; 99US-014294P.
 PR 28-JAN-2000; 2000US-0178717.
 XX
 XX (HUMA) HUMAN GENOME SCI INC.
 XX
 FI Ruben SM, Wei P, Ni J, Bastians GA, Shi Y;
 XX
 DR WPI: 2001-016507/02.
 XX
 PT Seven nucleic acid molecules encoding ADAM polypeptides
 PT disintegrin and metalloprotease domain, useful in the
 PT treatment and diagnosis of cancer, immune disorders, cell
 PT disorders and neurological diseases -
 PS
 PS Claim 11; Page 271-274; 287pp; English.
 XX
 CC The present invention relates to seven members of the ADAM
 CC which contain A Disintegrin And Metalloprotease domain;
 CC The ADAMs proteins and DNA may be used to treat disease,
 CC additive or preservative, for chromosome identification,
 CC for diagnosing a disorder related to the female reproductive
 CC particularly breast and/or ovary cancer. They are also
 CC gene therapy of breast and ovarian cancer.
 XX
 SQ Sequence 918 AA;
 Query Match 2.8%; Score 9; DB 23; Length 918;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 34 AGVLVAIIIV 42
 IIIIIIII
 DB 705 AGVLVAIIIV 713
 RESULT 26
 AAB24351 standard; protein; 918 AA.
 XX
 AC AAB24351;
 XX
 DT 04-OCT-2002 (first entry)
 XX
 DE Human novel protease #2.
 XX
 DE Human novel protease #2.
 KW Human, novel human protein; NHP; protease; melittin-beta
 KW metalloprotease; drug screening; clinical trial; meiotic;
 KW pharmacological; nutraceutical; gene therapy; chromoso
 OS Homo sapiens.
 XX
 PN W0200236759-A2.
 XX
 PD 10-MAY-2002.
 XX
 PF 30-OCT-2001; 2001WO-US48498.
 XX
 PR 01-NOV-2000; 2000US-244939P.
 XX
 XX (LEXI-) LEXICON GENETICS INC.
 XX
 PI Walke DM, Wilganowski NL, Friddle CJ;

XX WP1: 2002-479761/51.
 DR N-PSDB: AAD919150.
 XX
 PT New nucleic acid encoding a protein that shares sequence similarity
 PT with mammalian meltrin beta/ADAM 19 homology metalloproteinase for use in
 PT diagnosis and treatment, and in drug screening -
 XX
 ES Claim 1: Page 38-40; 48pp; English.
 XX
 CC The invention relates to novel human proteins (NHP), novel human
 CC proteases that shares sequence similarity with mammalian meltrin beta/
 CC ADAM 19 homologue metalloproteinase, and their corresponding nucleic acid
 CC sequences. NHP DNA is useful for diagnosis, drug screening, clinical
 CC trial monitoring, treatment of diseases and disorders and
 CC pharmacological, cosmetic and nutritional applications. It is also
 CC useful as hybridisation probes for screening libraries, and assessing
 CC gene expression patterns (particularly using a microarray or high-
 CC throughput chip format). It is also useful in restriction fragment
 CC length polymorphism (RFLP) analysis to identify specific individuals,
 CC and as a probe to identify the corresponding mutant allele in the
 CC libraries. It is also useful for the detection of mutant
 CC NHPs or inappropriately expressed NHPs for the diagnosis of disease.
 CC It is also useful in gene therapy. The present amino acid sequence
 CC is novel human protease. Human NHP gene is located at chromosome 5.
 CC
 XX Sequence 918 AA:
 SQ
 Query Match: 2.8%, Score 9, EB 23, Length 918,
 Best Local Similarity 100.0%, Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 AGVAVALLV 42
 ID 705 AGVAVALLV 713
 DB
 RESULT 27
 AAE19181
 ID AAE19181 standard, Protein, 918 AA.
 XX
 XX AAE19181:
 AC
 XX
 DT 21-MAY-2002 (first entry)
 DE
 XX Human protease, PRTS-18 protein.
 XX
 KW Human: protease; PRTS-18; enzyme; gastritis; cirrhosis; Crohn's disease;
 KW gastrointestinal disorder; autoimmune; inflammatory; cell proliferative;
 KW cardiovascular; developmental; epithelial; neurological; reproductive;
 KW AIDS; Acquired Immune Deficiency Syndrome; allergy; rheumatoid arthritis;
 KW anaemia; asthma; atelectasis; hypotension; myocardial infarction;
 KW hepatitis; cancer; psoriasis; Cushing's syndrome; hypothyroidism; eczema;
 KW epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;
 KW Pick's disease; infertility; vitiligo; drug screening; gene therapy.
 KW
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FH Peptide 1..22
 FT /label= Signal_peptide
 FT Protein 23..918
 FT /label= Mature_PRTS_18_protein
 FT Domain 700..721
 FT /label= Transmembrane_domain
 XX
 XX W:200208196 A2.
 XX
 PD 31-JAN-2002.
 XX
 XX 17-JUL-2001: 2001WC-0522397.
 XX
 XX 21-JUL-2000: 2000US-220063P
 PR

PR 28-JUL-2000: 2000US-221680P.
 PR 04 AUG 2000: 2000US-223544P.
 PR 11-AUG-2000: 2000US-224717P.
 PR 14-AUG-2000: 2000US-225988P.
 PR 24-AUG-2000: 2000US-227568P.
 XX
 PA (INCYTE GENOMICS INT.).
 XX
 PI Deleane AM, Gandhi AR, Hatfield AA, Lu DM, Patterson C;
 PI Tribouley CM, Das D, Kallick DA, Nguyen DB, Lee EA, Khan FA;
 PI Yue H, Au-Young J, Griffin JA, Policky JL, Ramkumar J, Yang J;
 PI Thangavelu K, Ding L, Kearney L, Baughn MR, Borowsky ML;
 PI Sanjanwalla MS, Yao MG, Burford N, Wallia NK, Lal P, Lee S, Todd S;
 PI Lo JP, Tang YL, Elliott VS, Azimzal Y, Lu Y;
 XX
 XX WP1: 2002-270062/25
 DR N-PSDB: AAD90585.
 XX
 PT New human protease polypeptide, useful in diagnosis, prevention and
 PT treatment of gastrointestinal, cardiovascular, autoimmune/inflammatory,
 PT cell proliferative, developmental, epithelial and neurological
 PT disorders
 XX
 ES Claim 1: Page 158-160; 182pp; English.
 PS
 XX
 CC The invention relates to an isolated human protease polypeptide (PRTS).
 CC PRTS protein and DNA are useful for diagnosing, treating and preventing
 CC gastrointestinal disorders (gastritis, cirrhosis, Crohn's disease),
 CC autoimmune/inflammatory disorders (AIDS, allergy, rheumatoid arthritis,
 CC anaemia, asthma), cardiovascular disorder (atherosclerosis, hypertension,
 CC myocardial infarction), cell proliferative disorders (hepatitis, cancer,
 CC psoriasis), developmental disorders (Cushing's syndrome, hypothyroidism),
 CC epithelial disorder (vitiligo, keloid, eczema), neurological disorders
 CC (epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease,
 CC Parkinson's disease), and reproductive disorders (infertility). PRTS
 CC protein is useful in a number of drug screening techniques and to
 CC analyse the proteome of a tissue or cell type. PRTS DNA is useful for
 CC creating knock-in humanised animals or transgenic animals to model human
 CC diseases. In somatic or germ-line gene therapy and in microarrays
 CC utilising fluids or tissues from patients to detect altered PKIN
 CC expression. The present sequence is human PRTS-18 protein.
 CC
 XX Sequence 918 AA:
 SQ
 Query Match: 2.8%, Score 9, EB 23, Length 918;
 Best Local Similarity 100.0%, Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 34 AGVAVALLV 42
 ID 705 AGVAVALLV 713
 DB
 RESULT 28
 AAE24350
 ID AAE24350 standard; Protein; 926 AA.
 XX
 XX AAE24350:
 AC
 XX
 DT 04 OCT 2002 (first entry)
 DE
 XX Human novel protease #1.
 XX
 KW Human: novel human protein; NHP; protease; meltrin-beta; ADAM 19;
 KW metalloproteinase; drug screening; clinical trial monitoring; cosmetic;
 KW pharmacological, nutritional, gene therapy, chromosome 5.
 KW
 XX Homo sapiens.
 OS
 XX
 XX W:200216759 A2.
 XX
 XX 10-MAY-2002.
 PD
 XX

CC proteases that shares sequence similarity with mammalian serine beta-
 CC ADAM 19 homologue metalloproteinase, and their corresponding antibody gene
 CC sequences. NHP DNA is useful for diagnosis, drug screening, clinical
 CC trial monitoring, treatment of diseases and disorders and
 CC pharmaceutical, cosmetic and nutritional applications. It is also
 CC useful as hybridisation probes for screening libraries, and assessing
 CC gene expression patterns (particularly using a microarray or high-
 CC throughput chip format). It is also useful in restriction fragment
 CC length polymorphism (RFLP) analysis to identify specific individuals,
 CC and as a probe to identify the corresponding mutant allele in the
 CC libraries. It is also useful for the detection of mutant
 CC NHPs or inappropriately expressed NHPs for the diagnosis of disease.
 CC It is also useful in gene therapy. The present amino acid sequence
 CC is novel human protease Human NHP gene is located at chromosome 5
 CC XX

SO Sequence 963 AA:

Query Match 2.8%; Score 9; DB 23; Length 963;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AGVIVAVIV 42
 |||||
 DB 705 AGVIVAVIV 713

RESULT 31
 AAV72965
 ID AAV72965 standard; peptide: 36 AA.
 AC AAV72965;
 XX
 DT 13-JUN-2001 (first entry)
 XX

DE Mouse type II integral membrane protein, CIRE-like fragment #17.
 XX
 KW Mouse: type II integral membrane protein; CIRE; anti-allergic; vaccine;
 KW immunosuppressive; autoimmune disorder; allergy; viral infection; APC;
 KW antigen presenting cell; bacterial infection; macrophage; dendritic cell;
 KW DC.
 XX
 OS Mus musculus.
 XX
 FN W0200119869-A1.
 XX
 PD 22-MAR-2001.
 XX
 PF 13-SEP-2000; 2000MO-AU01096.
 XX
 PE 13-SEP-1999; 99AU-0002789.
 XX
 PK 13-SEP-1999; 99AU-0002789.
 XX
 XX (COORD-) COINTEL QUEENSLAND INST MEDICAL RES.
 XX
 PA Caminschi L, Wright MD, Shortman KD;
 XX
 PI WPI: 2001-244788/25.
 DR N-PSDB; AAD03244.
 XX
 DR N-PSDB; AAD03244.
 XX

PT Novel type II integral membrane protein, CIRE, preferentially expressed
 PT in dendritic cells, macrophages and their precursors stimulatory to T
 PT cells, useful for identifying compounds having immunoregulatory
 PT activity -
 XX
 PS Disclosure: Page 27; 53pp; English.
 XX

CC The present sequence is mouse CIRE-like fragment derived from a protein
 CC expressed on dendritic cells which has similar function to CIRE protein.
 CC The invention relates to mouse type II integral membrane protein, CIRE,
 CC and its corresponding cDNA molecule which is preferentially expressed in
 CC dendritic cells, macrophages and their precursors stimulatory to T cells.
 CC CIRE is useful for screening immunoregulatory compounds and as a marker
 CC for dendritic cell (DC) population. CIRE specific ligand such as
 CC monoclonal antibodies are useful for isolating an antigen presenting cell

CC (Ab-2) from a biological sample. CIRE DNA is also used as tools to analyse
 CC the properties and functions of the CIRE gene/protein. For example, the
 CC CIRE DNA molecules are used to generate animal models, which lack functional
 CC CIRE genes and to isolate regulatory regions of the CIRE gene. Such
 CC regulatory regions are used to selectively express exogenous genes in DC
 CC or Ab-2. The invention also relates to methods for modulating an immune
 CC response in a subject by binding and inhibiting the function of an APC
 CC such as a myeloid DC. The CIRE-specific ligands are useful for modulating
 CC immune response by interfering with the function, migration or maturation
 CC of DC or APC. The ligands bind to and interfere with the function of
 CC myeloid dendritic cells such that antigen processing is undertaken by
 CC lymphoid cell which leads to immune suppression and thereby for treating
 CC allergies and autoimmune disorders. The ligands are also useful to block
 CC or enhance the interaction of viruses or bacteria with DC, for improving
 CC resistance to infection. These ligands are also used to target molecules
 CC such as vaccine components, to DC or APC.
 CC XX

SO Sequence 36 AA:

Query Match 2.5%; Score 8; DB 23; Length 36;
 Best Local Similarity 100.0%; Pred. No. 4.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 WMGSLDNLN 260
 |||||
 DB 15 WMGSLDNLN 22

RESULT 32
 AAV72966
 ID AAV72966 standard; peptide: 36 AA.
 AC AAV72966;
 XX
 DT 13-JUN-2001 (first entry)
 XX

DE Mouse type II integral membrane protein, CIRE-like fragment #18.
 XX
 KW Mouse: type II integral membrane protein; CIRE; anti-allergic; vaccine;
 KW immunosuppressive; autoimmune disorder; allergy; viral infection; APC;
 KW antigen presenting cell; bacterial infection; macrophage; dendritic cell;
 KW DC.
 XX
 OS Mus musculus.
 XX
 FN W0200119869-A1.
 XX
 PD 22-MAR-2001.
 XX
 PF 13-SEP-2000; 2000MO-AU01096.
 XX
 PE 13-SEP-1999; 99AU-0002789.
 XX
 PK 13-SEP-1999; 99AU-0002789.
 XX
 XX (COORD-) COINTEL QUEENSLAND INST MEDICAL RES.
 XX
 PA Caminschi L, Wright MD, Shortman KD;
 XX
 PI WPI: 2001-244788/25.
 DR N-PSDB; AAD03244.
 XX
 DR N-PSDB; AAD03244.
 XX

PT Novel type II integral membrane protein, CIRE, preferentially expressed
 PT in dendritic cells, macrophages and their precursors stimulatory to T
 PT cells, useful for identifying compounds having immunoregulatory
 PT activity -
 XX
 PS Disclosure: Page 27; 53pp; English.
 XX

CC The present sequence is mouse CIRE-like fragment derived from a protein
 CC expressed on dendritic cells which has similar function to CIRE protein.
 CC The invention relates to mouse type II integral membrane protein, CIRE,
 CC and its corresponding cDNA molecule which is preferentially expressed in
 CC dendritic cells, macrophages and their precursors stimulatory to T cells.
 CC CIRE is useful for screening immunoregulatory compounds and as a marker

PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132885.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134258.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137232.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139464.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140454.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 01-JUL-1999; 9905-0142154.
PR 02-JUL-1999; 9905-0142055.
PR 06-JUL-1999; 9905-0142390.
PR 08-JUL-1999; 9905-0142804.
PR 09-JUL-1999; 9905-0142920.
PR 12-JUL-1999; 9905-0143577.
PR 13-JUL-1999; 9905-0143624.
PR 14-JUL-1999; 9905-0143624.
PR 15-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 16-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.
PR 21-JUL-1999; 9905-0145086.
PR 21-JUL-1999; 9905-0145088.
PR 22-JUL-1999; 9905-0145085.
PR 22-JUL-1999; 9905-0145087.

PR 22-JUL-1999; 9905-0145089.
PR 22-JUL-1999; 9905-0145192.
PR 23-JUL-1999; 9905-0145145.
PR 23-JUL-1999; 9905-0145218.
PR 23-JUL-1999; 9905-0145224.
PR 26-JUL-1999; 9905-0145276.
PR 27-JUL-1999; 9905-0145913.
PR 27-JUL-1999; 9905-0145918.
PR 27-JUL-1999; 9905-0145919.
PR 28-JUL-1999; 9905-0145951.
PR 02-AUG-1999; 9905-0146386.
PR 02-AUG-1999; 9905-0146388.
PR 02-AUG-1999; 9905-0146389.
PR 03-AUG-1999; 9905-0147038.
PR 04-AUG-1999; 9905-0147204.
PR 04-AUG-1999; 9905-0147302.
PR 05-AUG-1999; 9905-0147192.
PR 05-AUG-1999; 9905-0147260.
PR 06-AUG-1999; 9905-0147304.
PR 06-AUG-1999; 9905-0147416.
PR 09-AUG-1999; 9905-0147493.
PR 09-AUG-1999; 9905-0147945.
PR 10-AUG-1999; 9905-0148171.
PR 11-AUG-1999; 9905-0148319.
PR 12-AUG-1999; 9905-0148341.
PR 13-AUG-1999; 9905-0148565.
PR 13-AUG-1999; 9905-0148684.
PR 16-AUG-1999; 9905-0149368.
PR 17-AUG-1999; 9905-0149175.
PR 18-AUG-1999; 9905-0149426.
PR 20-AUG-1999; 9905-0149724.
PR 20-AUG-1999; 9905-0149724.
PR 20-AUG-1999; 9905-0149929.
PR 23-AUG-1999; 9905-0149902.
PR 23-AUG-1999; 9905-0149930.
PR 25-AUG-1999; 9905-0150566.
PR 26-AUG-1999; 9905-0150884.
PR 27-AUG-1999; 9905-0151065.
PR 27-AUG-1999; 9905-0151066.
PR 27-AUG-1999; 9905-0151080.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0154070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.


```

PR 27-JUL-1999; 99US-0145914
PR 27-JUL-1999; 99US-0145914
PR 28-JUL-1999; 99US-0145951
PR 02-AUG-1999; 99US-0146386
PR 02-AUG-1999; 99US-0146388
PR 02-AUG-1999; 99US-0146384
PR 03-AUG-1999; 99US-0147038
PR 04-AUG-1999; 99US-0147204
PR 04-AUG-1999; 99US-0147302
PR 05-AUG-1999; 99US-0147192
PR 05-AUG-1999; 99US-0147260
PR 06-AUG-1999; 99US-0147404
PR 06-AUG-1999; 99US-0147416
PR 09-AUG-1999; 99US-0147493
PR 09-AUG-1999; 99US-0147935
PR 10-AUG-1999; 99US-0148177
PR 11-AUG-1999; 99US-0148319
PR 12-AUG-1999; 99US-0148341
PR 13-AUG-1999; 99US-0148565
PR 13-AUG-1999; 99US-0148684
PR 16-AUG-1999; 99US-0149368
PR 17-AUG-1999; 99US-0149175
PR 18-AUG-1999; 99US-0149426
PR 20-AUG-1999; 99US-0149722
PR 20-AUG-1999; 99US-0149723
PR 20-AUG-1999; 99US-0149929
PR 23-AUG-1999; 99US-0149932
PR 23-AUG-1999; 99US-0149940
PR 25-AUG-1999; 99US-0150566
PR 26-AUG-1999; 99US-0150884
PR 27-AUG-1999; 99US-0151065
PR 27-AUG-1999; 99US-0151066
PR 27-AUG-1999; 99US-0151080
PR 30-AUG-1999; 99US-0151303
PR 31-AUG-1999; 99US-0151438
PR 01-SEP-1999; 99US-0151930
PR 07-SEP-1999; 99US-0152363
PR 10-SEP-1999; 99US-0153070
PR 13-SEP-1999; 99US-0153358
PR 15-SEP-1999; 99US-0154018
PR 16-SEP-1999; 99US-0154039
PR 20-SEP-1999; 99US-0154779
PR 22-SEP-1999; 99US-0155139
PR 23-SEP-1999; 99US-0155486
PR 24-SEP-1999; 99US-0155659
PR 28-SEP-1999; 99US-0156458
PR 29-SEP-1999; 99US-0156596
PR 04-OCT-1999; 99US-0157117
PR 05-OCT-1999; 99US-0157553
PR 06-OCT-1999; 99US-0157865
PR 07-OCT-1999; 99US-0158029
PR 08-OCT-1999; 99US-0158232
PR 12-OCT-1999; 99US-0158369
PR 13-OCT-1999; 99US-0159293
PR 13-OCT-1999; 99US-0159294
PR 13-OCT-1999; 99US-0159295
PR 14-OCT-1999; 99US-0159329
PR 14-OCT-1999; 99US-0159330
PR 14-OCT-1999; 99US-0159331
PR 14-OCT-1999; 99US-0159637
PR 14-OCT-1999; 99US-0159638
PR 18-OCT-1999; 99US-0159584
PR 21-OCT-1999; 99US-0160741
PR 21-OCT-1999; 99US-0160767
PR 21-OCT-1999; 99US-0160768
PR 21-OCT-1999; 99US-0160770
PR 21-OCT-1999; 99US-0160814
PR 21-OCT-1999; 99US-0160815
PR 22-OCT-1999; 99US-0160980
PR 22-OCT-1999; 99US-0160981
PR 22-OCT-1999; 99US-0160984
PR 25-OCT-1999; 99US-0161404
PR 25-OCT-1999; 99US-0161405

```

```

PR 25-OCT-1999; 99US-0161406
PR 26-OCT-1999; 99US-0161459
PR 26-OCT-1999; 99US-0161360
PR 26-OCT-1999; 99US-0161361
PR 28-OCT-1999; 99US-0161920
PR 28-OCT-1999; 99US-0161922
PR 28-OCT-1999; 99US-0161993
PR 29-OCT-1999; 99US-0162142

Query Match
Best Local Similarity 2.5%; Score 8; DB 21; Length 213;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 PCWQWVD 269
DB 16 ESTWQWVD 24

RESULT 36
ID AAG78158 standard; Protein: 239 AA.
AC AAG78158;
XX 04-DEC-2001 (first entry)
XX
XX Homo sapiens.
XX
XX CN1300753-A.
XX
XX 27-JUN-2001.
XX
XX 22-DEC-1999; 99CN-0125674.
XX
XX 22-DEC-1999; 99CN-0125674.
XX
XX (BDB-) BDB: GENE DEV CO LTD SHANGHAI.
XX
XX Mao Y, Xie Y;
XX
XX WPI: 2001-530456/59.
XX
XX N-PSDB: AA164580.
XX
XX New polypeptide-human Epe protein 26 for treating diseases, such as,
XX cancer and human immunodeficiency virus infection.
XX
XX Claim 1: Page 24 25 (Disclosure); 31pp; Chinese.
XX
XX The invention relates to the human polypeptide human Epe protein 26, the
XX polynucleotide for coding it and use of the polypeptide in treating
XX diseases, including cancer, human immunodeficiency virus (HIV) infection.
XX
XX Sequence 239 AA:

Query Match
Best Local Similarity 2.5%; Score 8; DB 22; Length 239;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 QVSKVSS 50
DB 201 QVSKVSS 208

RESULT 37
ID AAG26290 standard; Protein: 279 AA.
XX
XX AAG26290;
AC AAG26290;

```



```

PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151067.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151940.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153788.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155553.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158332.
PR 12-OCT-1999; 990S-0158369.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159337.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160840.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 25-OCT-1999; 990S-0161454.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

```

```

Query Match: 2.5% Score 8; DB 21; Length 279;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 262 ESTWQWVD 269
DB 82 ESTWQWVD 89

RESULT 38
AA088024
ID AA088024 standard; Protein: 325 AA.
XX
AC AA088024;
XX
DI 05-JUN-2002 (first entry)
XX
DE Mouse protein encoded by Contig 1A.
XX
KW Nucleic acid library; immune response; asthma; COPD;
    airway hyperresponsiveness; bronchiolovular manifestation.

```

```

KW signature sequence; SS; chronic obstructive pulmonary disease;
KW allergic disease; rhinitis; atopic dermatitis; urticaria;
KW autoimmune disease; multiple sclerosis; inflammatory bowel disease;
KW allograft rejection; infectious disease.
XX calcium-activated chloride channel.
XX
XX Mus sp.
XX W200214366_A2.
XX
XX 21-FEB-2002.
XX
XX 16 AUG 2001; 2001W0-NL00610.
XX
XX 16 AUG 2000; 2000P-0202867.
XX
XX (UTUT-) RIKSINIV UTRECHT.
XX
XX Groot PC, Van Bergenhenegouwen BJ, Van Oosterhout ADM;
XX
XX WFL, 2000 211882790
XX
XX N PSDB, ABK47192.
XX
XX Nucleic acid library comprising genes which are capable of initiation,
XX proliferation and suppression of an immune response, especially an immune
XX response observed with airway hyper-responsiveness or asthma.
XX
XX Example 10, Fig 10; 120pp; English.
XX
XX The invention relates to a nucleic acid library comprising genes or
XX their fragments which are capable of modulating an immune response
XX observed with airway hyperresponsiveness and/or bronchiolovular
XX manifestations of asthma. Also included are a method for modulating an
XX immune response of an individual comprising modulating a gene comprising
XX a nucleic acid at least functionally equivalent to a nucleic acid
XX identifiable by a signature sequence (SS) given in the specification such
XX as R1-So-R1 A11, St01-A10, 3'02-1-G11, St01-A12, and R1-So-R1-B7, a
XX substance (for use as a nucleic acid) capable of modulating a gene
XX comprising a nucleic acid at least functionally equivalent to a nucleic
XX acid identifiable by SS and the use of a polypeptide substance derived
XX from a nucleic acid at least functionally equivalent to a nucleic acid
XX identifiable by SS for the production of an antagonist (for use as a
XX medicament) against the substance. The antagonist and substance are
XX useful for the treatment of an immune response observed with airway
XX hyper-responsiveness and/or bronchiolovular manifestations of asthma.
XX The method is useful for modulating the above immune response, where the
XX gene encodes a gene product capable of modulating the immune response.
XX The substance is useful for treating an immune response, particularly
XX asthma, chronic obstructive pulmonary disease (COPD), allergic diseases
XX (rhinitis, atopic dermatitis, urticaria), autoimmune diseases (e.g.
XX multiple sclerosis), inflammatory bowel disease, allograft rejection and
XX infectious disease. The present sequence is a mouse or human
XX protein encoded by a signature sequence gene or its homologue/function
XX equivalent.
XX
XX Sequence 125 AA;

```

```

Query Match: 2.5% Score 8; DB 23; Length 325;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 IVDGLISF 30
DB 54 IVDGLISF 61

RESULT 39
AA088026
ID AA088026 standard; Protein: 125 AA.
XX
AC AA088026;
XX
DI 05-JUN 2002 (first entry)
XX

```

XX Human DC-SIGN protein.
 DE
 XX
 XX Nucleic acid library: immune response; asthma; COPD;
 KM airway hyperresponsiveness; bronchoalveolar manifestation;
 KM signature sequence; SS: chronic obstructive pulmonary disease;
 KM allergic disease; rhinitis; atopic dermatitis; urticaria;
 KM allograft rejection; infectious disease; inflammatory bowel disease;
 KM calcium-activated chloride channel
 XX
 OS Homo sapiens.
 XX
 FN W0200214366-A2.
 XX
 PD 21-FEB-2002.
 XX
 PF 16-AUG-2001; 2001WO-NL00610.
 XX
 PK 16-AUG-2001; 2001EP-0202467
 XX
 PA (UUNT-) PIJXSUNIV UTRECHT.
 XX
 PI Groot PC, Van Bergenhenegouwen BJ, Van Oosterhout AJM;
 DR WPI: 2002-241888/29.
 XX
 PT Nucleic acid library comprising genes which are capable of initiation,
 PT progression and suppression of an immune response, especially an immune
 PT response observed with airway hyper-responsiveness of asthma
 XX
 PS Example 10; Fig 11; 120pp; English.
 XX
 CC The invention relates to a nucleic acid library comprising genes or
 CC their fragments which are capable of modulating an immune response
 CC observed with airway hyperresponsiveness and/or bronchoalveolar
 CC manifestations of asthma. Also included are a method for modulating an
 CC immune response of an individual comprising modulating a gene comprising
 CC a nucleic acid at least functionally equivalent to a nucleic acid
 CC identifiable by a signature sequence (SS) given in the specification such
 CC as RI-SO-RI-A11, SC01-A10, SV02-1-C11, SC01-A12, and RI-SO-RI-B7, a
 CC substance (for use as a medicament) capable of modulating a gene
 CC comprising a nucleic acid at least functionally equivalent to a nucleic
 CC acid identifiable by SS and the use of a proteaceous substance derived
 CC from a nucleic acid at least functionally equivalent to a nucleic acid
 CC identifiable by SS for the production of an antagonist (for use as a
 CC medicament) against the substance. The antagonist and substance are
 CC useful for the treatment of an immune response observed with asthma.
 CC hyperresponsiveness and/or bronchoalveolar manifestations of asthma.
 CC The method is useful for modulating the above immune response, where the
 CC gene encodes a gene product capable of modulating the immune response.
 CC The substance is useful for treating an immune response, particularly
 CC asthma, chronic obstructive pulmonary disease (COPD), allergic diseases
 CC (rhinitis, atopic dermatitis, urticaria), autoimmune diseases (e.g.
 CC multiple sclerosis), inflammatory bowel disease, allograft rejection and
 CC infectious disease. The present sequence is a mouse or human
 CC protein encoded by a signature sequence gene or its homologous/functional
 CC equivalent.
 XX
 XX
 SU Sequence 125 AA;
 XX
 Query Match 2.5%; Score 8; DR 23; Length 325;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 23 LVQLLSF 30
 |||||
 Db 54 LVQLLSF 61
 RESULT 40
 ABP25448
 ID ABP25448 standard; Protein: 356 AA.

XX
 AC ABP25448;
 XX
 XX 02-JUL-2002 (first entry)
 DT
 XX Streptococcus polypeptide SEQ ID No 72.
 DE
 XX Streptococcus; GAS; group A streptococcus; Streptococcus
 KM group A streptococcus; Streptococcus pyogenes; antibiotic
 KM antiinflammatory; infection; vaccine; meningitis; gene
 XX
 XX Streptococcus pyogenes.
 OS
 XX
 FN W0200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GH04789.
 XX
 PK 27-OCT-2001; 2001GB-0026344.
 XX
 PA 24-NOV-2000; 2000GB-0028747.
 XX
 PR 07-MAR-2001; 2001GB-0005649.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Masignani V, Margariti K, Grandi G,
 PI Telford J;
 DR WPI: 2002-352546/38.
 XX
 DR N-PSDR: ARN66074.
 XX
 PT New Streptococcus protein for the treatment or prevention of
 PT disease caused by Streptococcus bacteria, such as new
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3161; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895), it
 CC Streptococcus/GAS (Streptococcus agalactiae) or group A
 CC (Streptococcus pyogenes), comprising one of 5483 sequences
 CC the specification. The proteins have antibacterial and
 CC activity. (1), nucleic acids encoding (1), ARN66044-ARN713;
 CC antibodies that bind (1) are used in the manufacture of
 CC the treatment or prevention of infection or disease caused
 CC Streptococcus bacteria, particularly S. agalactiae and S.
 CC Nucleic acids encoding (1) are used to detect Streptococcus
 CC biological sample. (1) is used to determine whether a
 CC (1). A composition comprising (1) or a nucleic acid encod
 CC used as a vaccine or diagnostic composition. The disease
 CC Streptococcus that is prevented or treated may be mening
 CC acid encoding (1) may be used to recombinantly produce
 CC used in gene therapy. Antibodies to (1) are used for a
 CC chromatography, immunoassays, and distinguishment/identifi
 CC Streptococcus proteins.
 XX
 XX
 SU Sequence 356 AA;
 XX
 Query Match 2.5%; Score 8; DR 23; Length 37
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 8; Conservative 0; Mismatches 0; Indels
 QY 178 QELTDLKT 185
 |||||
 Db 78 QELTDLKT 85
 RESULT 41
 AA081511
 ID AA081511 standard; Protein: 472 AA.
 XX
 AC AA081511;
 XX

09-APR-2002 (first entry)
 C. elegans interflagellar transport protein, IFT52.
 Human, chlamydomonas, intraflagellar transport protein, Cht-2.
 anti-bacterial; protozoacide; contraceptive; anti-inflammatory; malaria;
 Chlamydomonas; IFT; mouse; kidney disease; retinal disorder; nematode;
 thyroid disorder; chondrocyte disease; olfactory disease; azospermia;
 primary ciliary dyskinesia, insect, protozoa, male contraceptive;
 parasitic disease; African sleeping sickness; trypanosomiasis;
 trichomoniasis; giardiasis; flagella.
 Caenorhabditis elegans.
 W0200190307-AZ.
 29-NOV-2001.
 24-MAY-2001; 2001WO-0517104.
 24-MAY-2000; 200003-2069234.
 (UYMA-) UNIV MASSACHUSETTS.
 Wilman GB, Pacour CJ, Rosenbaum JL, Cole DC.
 WPI: 2002 089926/12.
 Novel Chlamydomonas or mouse intraflagellar transport (IFT) protein or
 its variant, useful for identifying modulators that are useful for
 treating nematode, insect, protozoa or bacterial infection by
 inhibiting IFT.
 Disclosures: Figure 90; 132PF; English.
 The invention relates to an associated polypeptide (i) of chlamydomonas
 intraflagellar transport (IFT) particle protein, Cht-2, or mouse IFT
 particle protein 57. Also described are methods which are useful for:
 diagnosing kidney disease, retinal disorder, thyroid disorder,
 chondrocyte disease, olfactory disease, azospermia, or primary ciliary
 dyskinesia; for treating an infection in a mammal (a human) or plant
 caused by nematode, insect, protozoa or bacteria. The nucleic acid
 molecules are useful for diagnosis of disorders associated with aberrant
 expression of nucleic acid molecules and in genetic mapping and
 chromosome identification. The polypeptides are useful in generation of
 antibodies, as reagents in diagnostic assays, for the identification of
 other cellular gene products or compounds that can modulate the activity
 or expression of nucleic acids or polypeptides, and as pharmaceutical
 reagents useful for the treatment of disorders associated with aberrant
 expression or activity of the nucleic acids or polypeptides. IFT can
 serve as a male contraceptive. The IFT proteins can be targeted for
 treating parasitic diseases such as malaria, African sleeping sickness,
 trypanosomiasis, leishmaniasis, trichomoniasis, and giardiasis. An anti-IFT
 drug, taken orally, would inhibit assembly of flagella in newly divided
 Giardia and cause disassembly of previously formed flagella in non-
 dividing Giardia. An anti-IFT drug would block assembly of the flagellum
 and the flagellar sheath and affect the trypanosome's life cycle. The
 anti-IFT modulators can also be used to treat trichomonos vaginitis in
 humans. IFT inhibitors are also useful for combating phytopathogenic
 nematodes inhibiting insect pest by blocking sensory cilia function
 and/or assembly, leaving an insect unable to smell, unable to taste and
 unable to hear. AA081494-AA081529 represent intraflagellar transport
 particle amino acid sequences of the invention.

Query Match: 2.5%, Score 8, DB 23, Length 472;
 Best Local Similarity 100.0%, Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 169 DQSK0001 176
 1111111111
 21 DQSK0001 28

RESULT 42
 AA095815
 ID AA095815 standard; Protein: 596 AA.
 AA095815;
 26-JUN-2001 (first entry)
 Human protein sequence SWQ ID NO:18813.
 Human, primer, detection, diagnosis; antisense therapy; gene therapy.
 Homo sapiens.
 EP1074617-AZ.
 07-FEB-2001.
 28-JUL-2000; 2000EP-0116126.
 29-JUL-1999; 99JP-0248036.
 27-AUG-1999; 99JP-0300253.
 11-7AR-2000; 2000JP-0118776.
 02-MAY-2000; 2000JP-0183767.
 09-JUN-2000; 2000JP-0241899.
 (HELIX) HELIX RES INST.
 Ota T, Isachi T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 Ishii S, Shojiyama T, Wakamatsu A, Nagai K, Otsuki T;
 WPI: 2001-318749/34.
 First sets of synthesizing polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the
 full-length cDNAs.
 Claim 8, SWQ ID 18813, 2537PF; CD ROM; English.
 The present invention describes primer sets for synthesizing 5602
 full-length cDNAs defined in the specification. Where a primer set
 comprises (a) an oligo-dT primer and an oligonucleotide complementary
 to the complementary strand of a polynucleotide which comprises one of
 the 5602 nucleotide sequences defined in the specification, where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end
 sequence and an oligonucleotide comprising a sequence complementary to a
 polynucleotide which comprises a 3'-end sequence, where the
 oligonucleotide comprises at least 15 nucleotides and the combination of
 the 5'-end sequence/3'-end sequence is selected from those defined in
 the specification. The primer sets can be used in antisense therapy and
 in gene therapy. The primers are also useful for synthesizing polynucleotides,
 particularly full-length cDNAs. The primers are also useful for the
 detection and/or diagnosis of the abnormality of the proteins encoded by
 the full-length cDNAs. The primers allow obtaining of the full-length
 cDNAs easily without any specialised methods. AA003146 to AA013628 and
 AA013633 to AA018742 represent human cDNA sequences; AA092446 to
 AA096892 represent human amino acid sequences; and AA013632
 represent oligonucleotides, all of which are used in the exemplification
 of the present invention.

Query Match: 2.5%, Score 8, DB 22, Length 596;
 Best Local Similarity 100.0%, Pred. No. 65;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 43 QVSKRPSS 50
 1111111111

Db 558 QVSKVPSS 565

RESULT 43

AAV24094 standard; Protein: 763 AA.

AAV24094:

04-SEP-1999 (first entry)

Mouse Epc1.

Epc1, Epc2, enhancer of polycomb, tumour suppressor, diagnosis; proliferative disorder; developmental disorder; neoplasia.

Mus sp.

W0943986-A1.

08-JUL-1999.

28-DEC-1998: 98W0-0527667.

10-DEC-1997: 97US-0068957.

(CHIR) CHIRON CORP.

Randazzo F:

WPI: 1999-419108/35.

N-PSDB: AAX83993.

Mammalian tumour suppressor, termed enhancer of polycomb, useful for therapy, prognosis and diagnosis of proliferative and developmental disorders

Claim 2: Page 49; 55pp: English.

The present sequence represents a mouse tumour suppressor, designated enhancer of polycomb (Epc). Human Epc genes can be used to identify a 10-11 region of human chromosome 10 and a q13.3 region of human chromosome 22. The Epc genes and expression products are useful as therapeutic, prognostic and diagnostic tools for proliferative and developmental disorders, e.g. neoplasia.

Sequence 763 AA:

Query Match 2.58; Score 8; Db 20; Length 763;
Best Local Similarity 100.0%; Pred. No. 81;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 QVSKVPSS 50

Db 725 QVSKVPSS 732

RESULT 44

AAH45133 standard; Protein: 27 AA.

AAH45133:

12-FEB-2001 (first entry)

Human secreted protein sequence encoded by gene 14 SFG in No-74

Human; secreted protein; immunosuppressive; antiarthritic; arthritic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neurotrophic; neuroprotective; antibacterial; virucide; fungicide; cancer; optimaological; autoimmune disease; hyperproliferative disorder; cardiovascular disorder; cerebrovascular disorder; wound healing; nervous system disorder; aging; chemotaxis.

XX Homo sapiens.

W0200058467-A1.

05-OCT-2000.

24-MAR-2000: 2000W0-0507505.

26-MAR-1999: 99US-0126502.

17-DEC-1999: 99US-0172410.

(HUMA-) HUMAN GENOME SCI INT.

Rosen CA, Ruben SM, Komatsoulis G:

WPI: 2000-611712/58.

N-PSDB: AAC80544.

Nucleic acid molecules encoding human secreted proteins, preventing, treating or ameliorating a disorder, e.g. Al

Parkinson's diseases and cancers

Claim 11: Page 385; 440pp: English.

Polynucleotide sequences AAC80544-C80580 represent cDNA secreted proteins AAB45120-145169. Sequences AAB45170-145179 alternative polypeptides encoded by the genes, and amino acids to which they are homologous. The genes and proteins have dependent on the tissues and cells in which they are expressed. Their activities include immunosuppressive; antiarthritic; antineoplastic; antiproliferative; cytostatic; antiarthritic; cerebroprotective; neurotrophic; neuroprotective; antibacterial; fungicide; and optimaological. The secreted proteins, if antagonists and agonists may be useful in treating, preventing, diagnosing diseases and disorders such as autoimmune disease, rheumatoid arthritis, hyperproliferative disorders, e.g. the breast or liver, cardiovascular disorders e.g. coronary artery disease, Alzheimer's disease, infectious diseases, system disorders e.g. AIDS, bacterial, viral and fungal diseases e.g. AIDS. The polypeptides can also be used to aid wound healing, cell proliferation, to prevent skin aging due to sunburn, organs before transplantation, for supporting cell culture, tissues, to regenerate tissues and in chemotaxis. The proteins also be used as a food additive or preservative to improve storage capabilities. AAC80522-C80530 and AAB45119 represent the invention.

Sequence 27 AA:

Query Match 2.2%; Score 7; Db 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 34;

Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 32 LLAGVIV 38

Db 20 LLAGVIV 26

RESULT 45

AAR22106 standard; Protein: 29 AA.

AAR22106:

15-JUL-1992 (first entry)

Peptide 3-29 derived from CD18 leukocyte integrin.

Cell adhesion; intercellular adhesion molecule; endothelial inflammation; leukocyte chemotaxis; rhinovirus; common

```

XX W9204473-A.
PN
XX
XX 05-MAR-1992.
FD
XX 24-AUG-1991; 91WO-0006063.
PF
XX 27-AUG-1990; 90US-0573624.
PR
XX
XX (CETU ) CETUS CORP.
PA
XX
XX Liu DY, Kaymakalan Z, Mundy K:
PI
XX WPI; 1992-096833/12.
DR
XX
XX Peptide(s) derived from beta sub-unit CD18 of leukocyte integrins
PT - prevent leukocyte binding to ICAM and leukocyte chemotaxis, for
PI treating inflammatory diseases and rhinoviral infection
XX
XX claim 2; Page 23; 31pp; English.
PS
XX
XX This peptide (derived from CD18 see AAR24256) was synthesised and
CC tested for capacity to interfere with or block adhesion of
CC polymorphonuclear leukocytes to human endothelial cell monolayers.
CC The peptide shows inhibitory activity at 10(power -4)M. The peptide
CC can be used to treat/prevent inflammation and can inhibit binding
CC of viruses to endothelium. See AAQ22780 and AAR22104 R22112.
XX
XX
XX Sequence 29 AA:
SU

```

```

Query Match          2.2%; Score 7; DB 13; Length 29;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

UY 71 AVGEISE 77
   111111
Db 5 AVGEISE 11

```

Search completed: December 7, 2002, 11:15:26
Job time : 40 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM protein - protein search, using sw model

Run on: December 7, 2002, 11:14:40 : Search time 17 seconds
(without alignments)
562.497 Million cell updates/sec

Title: US-09-831-458A-12

Sequence: 1 MDSKEPRVQGLIAGLGH... RGVNDNWICKKPAACEPDE 325

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 5

Total number of hits satisfying chosen parameters: 6756

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA:

- 1: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/1/aa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/1/aa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/1/aa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/1/aa/PCITUS_COMB.pep:*
- 6: /cgn2_6/prodata/1/aa/Quicklial.pep:*

Pred. No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	168	51.7	296	4	US-09-517-605-17 Sequence 17, Appl
2	78	24.0	494	4	US-09-517-605-2 Sequence 2, Appl
3	9	2.8	751	2	US-08-836-443-3 Sequence 3, Appl
4	7	2.2	154	1	US-08-150-945-67 Sequence 67, Appl
5	7	2.2	154	4	US-08-976-161-67 Sequence 67, Appl
6	7	2.2	252	5	PCT-US96-01314-58 Sequence 58, Appl
7	7	2.2	253	5	PCT-US96-01314-53 Sequence 53, Appl
8	7	2.2	548	2	US-09-932-312-4 Sequence 4, Appl
9	7	2.2	548	2	US-08-993-318A-4 Sequence 4, Appl
10	7	2.2	548	4	US-09-199-886-4 Sequence 4, Appl
11	7	2.2	548	4	US-09-396-260-4 Sequence 4, Appl
12	7	2.2	548	4	US-09-576-281-4 Sequence 4, Appl
13	7	2.2	676	3	US-08-630-172-10 Sequence 10, Appl
14	7	2.2	676	4	US-09-375-419-10 Sequence 10, Appl
15	7	2.2	720	3	US-09-257-799-48 Sequence 48, Appl
16	7	2.2	720	3	US-08-920-919A-48 Sequence 48, Appl
17	7	2.2	769	2	US-08-789-078-1 Sequence 1, Appl
18	7	2.2	769	2	US-08-752-633-1 Sequence 1, Appl
19	7	2.2	769	2	US-08-476-062A-45 Sequence 45, Appl
20	7	2.2	769	2	US-07-728-215-31 Sequence 31, Appl
21	7	2.2	769	4	US-08-938-085A-31 Sequence 31, Appl
22	7	2.2	769	5	PCT-US95-04886-1 Sequence 45, Appl
23	7	2.2	769	5	PCT-US95-04886-1 Sequence 45, Appl
24	6	1.8	8	6	US-09-5185431-19 Patent No. 5185431
25	6	1.8	9	6	US-08-480-226-12 Patent No. 5185431
26	6	1.8	14	2	US-08-659-235C-12 Patent No. 5185431
27	6	1.8	14	2	US-08-659-235C-12 Patent No. 5185431

28	6	1.8	19	1	US-07-908-417-10	App1
29	6	1.8	19	5	PCT-US94-09171-19	App1
30	6	1.5	25	4	US-08-819-624-24	App1
31	6	1.8	25	4	US-09-150-812-24	App1
32	6	1.8	33	4	US-08-849-624-23	App1
33	6	1.8	33	4	US-09-150-812-23	App1
34	6	1.8	37	4	US-08-849-624-25	App1
35	6	1.8	37	4	US-09-150-812-25	App1
36	6	1.8	39	1	US-08-189-431-143	App1
37	6	1.8	39	2	US-08-471-066-143	App1
38	6	1.8	40	4	US-08-689-421-1	App1
39	6	1.8	40	4	US-09-489-528-1	App1
40	6	1.8	40	4	US-09-181-827A-1	App1
41	6	1.8	45	4	US-08-849-624-26	App1
42	6	1.8	49	4	US-09-150-812-26	App1
43	6	1.8	52	4	US-08-946-165A-408	App1
44	6	1.8	58	1	US-07-664-989B-78	App1
45	6	1.8	58	1	US-08-458-160-55	App1

ALIGNMENTS

```

RESULT 1
US-09-517-605-17
Sequence 17, Application US/09/17605
Patent No. 6391567
GENERAL INFORMATION:
APPLICANT: Liltman, Dan R.
APPLICANT: Kwon, Douglas S.
APPLICANT: van Kooyk, Yvette
APPLICANT: Geljtenbeck, Ineo
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF REPE
FILE REFERENCE: 1049-1-017
CURRENT AFFILIATION NUMBER: US/09/517,605
CURRENT FILING DATE: 2000-03-02
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patentio Ver. 2.0
SEQ ID NO 17
LENGTH: 296
TYPE: PRT
ORGANISM: Homo sapiens
US-09-517-605-17

Query Match      51.7% Score 168 DB 4 Length
Best Local Similarity 100.0% Pred. No. 7.5e-154
Matches 168, Conservative 0, Mismatches 0, Indels 0

QY 1 MDSKEPRVQGLIAGLGHAVLQIISPMLAGYVALLVQSVSPSS
Db 1 MDSKEPRVQGLIAGLGHAVLQIISPMLAGYVALLVQSVSPSS
QY 61 IYQNLIDLKAAGVETSEKSLQIYQELTQKAAGVETPEKSKLQIYQ
Db 61 IYQNLIDLKAAGVETSEKSLQIYQELTQKAAGVETPEKSKLQIYQ
QY 121 IPRSKQIYQIYQPTTPIKAAGVETPEKSKLQIYQELTQKAAGVET
Db 121 IPRSKQIYQIYQPTTPIKAAGVETPEKSKLQIYQELTQKAAGVET
QY 121 IPRSKQIYQIYQPTTPIKAAGVETPEKSKLQIYQELTQKAAGVET
Db 121 IPRSKQIYQIYQPTTPIKAAGVETPEKSKLQIYQELTQKAAGVET

RESULT 2
US-09-517-605-2
Sequence 2, Application US/09/17605
Patent No. 6391567
GENERAL INFORMATION:
APPLICANT: Liltman, Dan R.
APPLICANT: Kwon, Douglas S.
APPLICANT: van Kooyk, Yvette
APPLICANT: Geljtenbeck, Ineo
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF REPE

```

```

FILE REFERENCE: 1049-1-017
CURRENT APPLICATION NUMBER: US/09/517,605
CURRENT FILING DATE: 2000-03-02
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 2
LENGTH: 404
TYPE: PRT
ORGANISM: Homo sapiens
US-09-517-605-2

```

```

Query Match          24.0%; Score 78; DB 4; Length 404;
Best Local Similarity 100.0%; Pred. No. 4, 1a-b7;
Matches /8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 58 QDAIYQNTQQLKAAGVLSKSKSLDEIYQELTQQLKAAGVLPKSKLOEIYQELTPKAA 117
   |||
Db 74 QDAIYQNTQQLKAAGVLSKSKSLDEIYQELTQQLKAAGVLPKSKLOEIYQELTPKAA 114
   |||
QY 118 VGEIPEKSKLOEIYQELT 135
   |||
Db 134 VGEIPEKSKLOEIYQELT 151

```

```

RESULT 3
US-08-836-443-3
Sequence 3, Application US/08836443
Patent No. 5883241
GENERAL INFORMATION:
APPLICANT: DOHERTY, Andrew, J.P
APPLICANT: SLOCOMBE, Patrick, M.
TITLE OF INVENTION: DNA SEQUENCES CODING FOR A
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: PastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,443
FILING DATE: 01-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9704420
FILING DATE: 05-SEP-1997
APPLICATION NUMBER: 9612145, 4
FILING DATE: 11-JUN-1996
APPLICATION NUMBER: 9526230, 9
FILING DATE: 21-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
REFERENCE/DOCKET NUMBER: 47424
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 751 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLAGY: linear
MOLECULE TYPE: NO. 5883241e
US-08-836-443-3

```

```

Query Match          2.8%; Score 3; DB 2; Length 751;
Best Local Similarity 100.0%; Pred. No. 1, 9;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 34 AGYVALIV 42
   |||
Db 540 AGYVALIV 538

```

```

RESULT 4
US-08-450-945-67
Sequence 67, Application US/08450945
Patent No. 5783383
GENERAL INFORMATION:
APPLICANT: Kondo, Kazuhito
APPLICANT: Morasaki, Edward S. Jr.
TITLE OF INVENTION: LATENT TRANSPLANTS AND PROMOTERS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESS: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,945
FILING DATE: 23-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Shultz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 154 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-450-945-67
Query Match          2.2%; Score 7; DB 1; Length 154;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 35 GVLVALIV 41
   |||
Db 23 GVLVALIV 29

```

```

RESULT 5
US-08-976-161-67
Sequence 67, Application US/08976161
Patent No. 614542
GENERAL INFORMATION:
APPLICANT: Kondo, Kazuhito
APPLICANT: Morasaki, Edward S. Jr.
TITLE OF INVENTION: LATENT TRANSPLANTS AND PROMOTERS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESS: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250

```

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentia Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08/976,161
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/450,945
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 67:
SEQUENCE CHARACTERISTICS:
LENGTH: 154 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-976-161-67

Query Match 2.28; Score 7; DB 4; Length 154;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 35 GYLVAII.41
IDB 23 GYLVAII.29

RESULT 6
PCT-US96-01314-58
Sequence 58, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/380,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906

TELEX: 200154
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 252 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01314-58

Query Match 2.28; Score 7; DB 5; Length 252;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 71 AVGELSE 77
IDB 232 AVGELSE 238

RESULT 7
PCT-US96-01314-53
Sequence 53, Application PC/TUS9601314
GENERAL INFORMATION:
APPLICANT: M. Amin Arnaout
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
NUMBER OF SEQUENCES: 78
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
REFERENCE/DOCKET NUMBER: 00786/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 253 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-01314-53

Query Match 2.28; Score 7; DB 5; Length 253;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0;
QY 71 AVGELSE 77
IDB 233 AVGELSE 239

RESULT 8
US-09-032-315-4

Sequence 4, Application US/09032315
Patent No. 5985818
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Xu, Feng
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 598818th, No. 598818th, America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/032,315
FILING DATE: 27-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,943
REFERENCE/DOCKET NUMBER: 5200,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-032-315-4

Query Match 2.2% Score 71 DB 2, Length 548,
Best Local Similarity 100.0%, Freq. No. 1.2e+02,
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

QY 308 DWDNYWI 314
Db 275 DWDNYWI 281

RESULT 9
US-08-993-318A-4
Sequence 4, Application US/08993318A
Patent No. 5998353
GENERAL INFORMATION:
APPLICANT: Pedersen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Rasmussen, Gertel
APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5998353th, No. 5998353th, America
STREET: 405 Lexington Avenue
CITY: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/993,318A

FILING DATE: December 18, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeria A.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5032,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-993-318A-4

Query Match 2.2% Score 71 DB 2, Length 548,
Best Local Similarity 100.0%, Freq. No. 1.2e+02,
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

QY 308 DWDNYWI 314
Db 275 DWDNYWI 281

RESULT 10
US-09-399-886-4
Sequence 4, Application US/09399886
Patent No. 6140092
GENERAL INFORMATION:
APPLICANT: Pedersen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Rasmussen, Gertel
APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6140092th, No. 6140092th, America
STREET: 405 Lexington Avenue
CITY: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/399,886
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,318
FILING DATE: December 18, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeria A.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 5032,200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-344-886-4

Query Match
Best Local Similarity 100.0%; Score 7, DB 4, Length 548,
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 408 DVDNYMI 314
DB 275 DVDNYMI 281

RESULT 11

US-09-346-260-4
Sequence 4, Application US/09396260
Patent No. 6184015
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Xu, Feng
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6184015, No. 6184015, 1st of No. 6184015, America, Inc.,
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/346,260
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/032,315
FILING DATE: 27-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 5200, 200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-346-260-4

Query Match
Best Local Similarity 100.0%; Score 7, DB 4, Length 548,
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 408 DVDNYMI 314
DB 275 DVDNYMI 281

RESULT 12

US-09-576-281-4
Sequence 4, Application US/09576281
Patent No. 6277611
GENERAL INFORMATION:
APPLICANT: Peperisen, Anders
APPLICANT: Svendsen, Allan
APPLICANT: Schneider, Palle
APPLICANT: Kasmussen, Grethe

APPLICANT: Cherry, Joel
TITLE OF INVENTION: LACCASE MUTANTS
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6277611, No. 6277611, 1st of No. 6277611,
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (Free)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/576,281
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/344,418
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valetta A.
REGISTRATION NUMBER: 34,728
REFERENCE/DOCKET NUMBER: 6042, 200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0124
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 548 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-576-281-4

Query Match
Best Local Similarity 100.0%; Score 7, DB 4, Length 548,
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 308 DVDNYMI 314
DB 275 DVDNYMI 281

RESULT 13

US-08-630-172-10
Sequence 10, Application US/08630172
Patent No. 6060054
GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PROTECTANT AND PROCESS FOR IT
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 45th Floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/630,172
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.
PRIOR APPLICATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 676 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-630-172-10

Query Match: 2.28; Score 7; DB 3; Length 676;
Best Local Similarity 100.0%; Prod No 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 71 AVECLSE 77
|||||
DB 319 AVECLSE 325

RESULT 14
US-09-375-419-10
Sequence 10, Application US/09375419
Patent No. 6264950
GENERAL INFORMATION:
APPLICANT: Staerz, Uwe
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T
TITLE OF INVENTION: LYMPHOCYTE VETO
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, 35th floor
CITY: Denver
STATE: Colorado
COUNTRY: U.S.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/375,419
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,172
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
PRIOR APPLICATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-36
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 676 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-475-419-10

Query Match: 2.28; Score 7; DB 4; Length 676;
Best Local Similarity 100.0%; Prod No 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 71 AVECLSE 77
|||||
DB 319 AVECLSE 325

RESULT 15
US-09-257-799-48
Sequence 48, Application US/09257799
Patent No. 6020178
GENERAL INFORMATION:
APPLICANT: Aschlimann, Daniel
TITLE OF INVENTION: Transglutaminase and Gene Encoding Same
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/257,799
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/920,919
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J.
PRIOR APPLICATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296, 94509
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-257-799-48

Query Match: 2.28; Score 7; DB 3; Length 720;
Best Local Similarity 100.0%; Prod No 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 269 DCSFLSP 275
|||||
DB 551 DCSFLSP 557

RESULT 16
US-08-920-919A-48
Sequence 48, Application US/08920919A
Patent No. 6114119
GENERAL INFORMATION:
APPLICANT: Aschlimann, Daniel
TITLE OF INVENTION: Transglutaminase and Gene Encoding Same
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,919A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 960296, 94509
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 720 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-920-919A-48

Query Match 2.28; Score 7; DB 3; Length 720;
Best Local Similarity 100.0%; Pred. No. 1.5e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 269 DGSPLSP 275
DB 551 DGSPLSP 557

RESULT 17
US-08-789-078-1
Sequence 1, Application US/08789078
Patent No. 5843885
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H.
APPLICANT: Slahaan, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbets, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/789,078
FILING DATE: 03-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/229,513
FILING DATE: 19-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26242
REFERENCE/DOCKET NUMBER: 22833
TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9050
TELEFAX: 816/474-9057
TELEX: 434-363
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid

TOPOLOGY: both
MOLECULE TYPE: protein
HYPOTHEICAL: N
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: tonsil
FEATURE:
NAME/KEY: Region
LOCATION: 1..22
OTHER INFORMATION: /label "signal"
OTHER INFORMATION: /note "signal sequence"
FEATURE:
NAME/KEY: Region
LOCATION: 449..496
OTHER INFORMATION: /label "repeat"
OTHER INFORMATION: /note "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 497..540
OTHER INFORMATION: /label "repeat"
OTHER INFORMATION: /note "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 541..581
OTHER INFORMATION: /label "repeat"
OTHER INFORMATION: /note "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 582..617
OTHER INFORMATION: /label "repeat"
OTHER INFORMATION: /note "cysteine rich repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 701..723
OTHER INFORMATION: /label "transmembrane domain"
OTHER INFORMATION: /note "transmembrane domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 724..769
OTHER INFORMATION: /label "cyto"
OTHER INFORMATION: /note "cytoplasmic domain"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
TIBBETS, Power,
TITLE: LFA-1 Amino acid sequence (K2) (from human
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 96-96
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
US-08-789-078-1
Query Match 2.28; Score 7; DB 2; Length 7
Best Local Similarity 100.0%; Pred. No. 1.5e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 71 AVGLSP 77
DB 341 AVGLSP 347

RESULT 18
US-08-752-633-1
Sequence 1, Application US/08752633
Patent No. 5863889
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H.
APPLICANT: Slahaan, Teruna
APPLICANT: Chan, Marcia
APPLICANT: Tibbets, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INDUCING
NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:
 ADDRESSEE: HOVEY, WILLIAMS, TIMMONS & COLLINS
 STREET: 1101 Walnut St.
 CITY: Kansas City
 STATE: MO
 COUNTRY: USA
 ZIP: 64106
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1 0, Version #1 25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/752,633
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Collins, John M.
 REGISTRATION NUMBER: 36262
 REFERENCE/DOCKET NUMBER: 22833
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (816)474-9050
 TELEFAX: (816)474-9057
 TELEX: 434-363
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 TOPOLOGY: both
 MOLECULE TYPE: protein
 HYPOTHEetical: NO
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens
 TISSUE TYPE: Tonsil
 FEATURE:
 NAME/KEY: Region
 LOCATION: 1..22
 OTHER INFORMATION: /label= "signal sequence"
 FEATURE:
 NAME/KEY: Region
 LOCATION: 449..496
 OTHER INFORMATION: /label= "repeat"
 FEATURE:
 NAME/KEY: Region
 LOCATION: 497..540
 OTHER INFORMATION: /label= "cysteine rich repeat"
 FEATURE:
 NAME/KEY: Region
 LOCATION: 541..581
 OTHER INFORMATION: /label= "repeat"
 OTHER INFORMATION: /note= "cysteine rich repeat"
 FEATURE:
 NAME/KEY: Region
 LOCATION: 582..617
 OTHER INFORMATION: /label= "repeat"
 OTHER INFORMATION: /note= "cysteine rich repeat"
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 701..723
 OTHER INFORMATION: /label= "transmembrane domain"
 OTHER INFORMATION: /note= "transmembrane domain"
 FEATURE:
 NAME/KEY: Domain
 LOCATION: 724..769
 OTHER INFORMATION: /label= "cyto"
 OTHER INFORMATION: /note= "cytoplasmic domain"
 PUBLICATION INFORMATION:
 AUTHORS: Pigott,
 TITLE: LFA-1 Amino acid sequence (B2) (from human

TITLE: Tonsil
 JOURNAL: The Adhesion Molecule Facts Book
 PAGES: 96-96
 DATE: 1993
 RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
 US-08-752-633-1
 Query Match: 2.2% Score 7, DB 2, Length 769;
 Best local similarity: 100.0%, Pred. No. 1,66-02;
 Matches 7, conservative 0, Mismatches 0, Indels 0;
 Db 341 AVERAGE 447
 CQ 71 AVERAGE 77
 11111111
 RESULT 19
 US-08-476-062A-45
 Sequence 47, Application US/02176062A
 Patent No. 5877275
 GENERAL INFORMATION:
 APPLICANT: Arnaud, M. Amin
 TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY
 TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS
 NUMBER OF SEQUENCES: 53
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: Windows95
 SOFTWARE: FASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476,062A
 FILING DATE: 07-JUN-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/216,081
 FILING DATE: 21-MAR-1994
 APPLICATION NUMBER: 07/647,840
 FILING DATE: 04-JAN-1991
 APPLICATION NUMBER: 07/549,842
 FILING DATE: 18-JUN-1990
 APPLICATION NUMBER: 07/212,573
 FILING DATE: 28-JUN-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Freeman, John W.
 REGISTRATION NUMBER: 29,066
 REFERENCE/DOCKET NUMBER: 08785/068003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 45:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 769 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-476-062A-45
 Query Match: 2.2% Score 7, DB 2, Length 769;
 Best local similarity: 100.0%, Pred. No. 1,66-02;
 Matches 7, conservative 0, Mismatches 0, Indels 0;
 Db 341 AVERAGE 447

RESULT 20
US-07-728-215-31
Sequence 31, Application US/07728215
Patent No. 5962643
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
TITLE OF INVENTION: A No. 5962643e1 Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/728,215
FILING DATE: 19910711
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-07-728-215-31
Query Match 2.2% Score 7: DB 2: Length 769:
Best Local Similarity 100.0%: Pred. No. 1.6e+02:
Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
QY 71 ANCELS 77
DB 341 ANCELS 347

RESULT 21
US-08-938-085A-31
Sequence 31, Application US/08938085A
Patent No. 6339148
GENERAL INFORMATION:
APPLICANT: Sheppard, Dean
APPLICANT: Quaranta, Vito
TITLE OF INVENTION: A No. 6339148e1 Integrin Beta Subunit and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 85204556.085A
FILING DATE: 26-SEP-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,215
FILING DATE: 11-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parent, Annette S.
REGISTRATION NUMBER: 42,058
REFERENCE/DOCKET NUMBER: 02407-08021005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-938-085A-31
Query Match 2.2% Score 7: DB 4: Length 76:
Best Local Similarity 100.0%: Pred. No. 1.6e+02:
Matches 7: Conservative 0: Mismatches 0: Indels 0:
QY 71 ANCELS 77
DB 341 ANCELS 347

RESULT 22
PCT-US95-04886-1
Sequence 1, Application PC/TUS9504886
GENERAL INFORMATION:
APPLICANT: Benedict, Stephen H.
APPLICANT: Siahaan, Teruua
APPLICANT: Chan, Marcia
APPLICANT: Tibbels, Scott
TITLE OF INVENTION: ICAM-1/LFA-1 PEPTIDES FOR INJECTION
TITLE OF INVENTION: IMMUNE TOLERANCE
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOVEY, WILLIAMS, LIMMONS & COLLINS
STREET: 1101 Walnut St.
CITY: Kansas City
STATE: MO
COUNTRY: USA
ZIP: 64106
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/TUS95/04886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Collins, John M.
REGISTRATION NUMBER: 26472
REFERENCE/DOCKET NUMBER: 22844
TELECOMMUNICATION INFORMATION:
TELEPHONE: (816)474-9057
TELEFAX: 434-363
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids

```

TYPE: amino acid
TOPOLOGY: both
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: tonsil
FEATURE:
NAME/KEY: Region
LOCATION: 1..22
OTHER INFORMATION: /label= signal
OTHER INFORMATION: /note= "signal sequence"
FEATURE:
NAME/KEY: Region
LOCATION: 449..496
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 497..540
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 541..581
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Region
LOCATION: 582..617
OTHER INFORMATION: /label= repeat
OTHER INFORMATION: /note= "cysteine rich repeat"
FEATURE:
NAME/KEY: Domain
LOCATION: 701..723
OTHER INFORMATION: /label= trans
OTHER INFORMATION: /note= "transmembrane domain"
FEATURE:
NAME/KEY: Domain
LOCATION: 724..769
OTHER INFORMATION: /label= cyto
OTHER INFORMATION: /note= "cytoplasmic domain"
PUBLICATION INFORMATION:
AUTHORS: Pigott,
TITLE: LFA-1 Amino acid sequence (B2) (From human
TITLE: tonsil)
JOURNAL: The Adhesion Molecule Facts Book
PAGES: 96-96
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 769
PCT-US96-04886-1

Query Match          2.2%; Score 7; DB 5; Length 769;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 71 ANGEISE 77
DB 341 ANGEISE 347

```

```

CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/01314
FILING DATE: 30-JAN-96
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,167
FILING DATE: 30-JAN-95
ATTORNEY/AGENT INFORMATION:
NAME: John W. Freeman
REGISTRATION NUMBER: 29,066
PREFERENCE/RIGHT NUMBER: 00784/267001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 769
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US96-01314-45

Query Match          2.2%; Score 7; DB 5; Length 769;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 71 ANGEISE 77
DB 341 ANGEISE 347

RESULT 24
5185431-19
Patent No. 5185431
APPLICANT: YOSHIMATSU, KENTARO;SHIKATA, YASUSHI;TANAKA, ISAO;
HASEGAWA, YOSHIKAZU;SETO, TOSHIO;OSAWA, TOSHIO
TITLE OF INVENTION: RECOMBINANT NATURAL KILLER CELL ACTIVATOR
NUMBER OF SEQUENCES: 31
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/07/392,841
FILING DATE: 11-AUG-1989
SEQ ID NO:19
LENGTH: 8
5185431-19

Query Match          1.8%; Score 6; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 266 QWVDGS 271
DB 2 QWVDGS 7

RESULT 25
5185431-20
Patent No. 5185431
APPLICANT: YOSHIMATSU, KENTARO;SHIKATA, YASUSHI;TANAKA, ISAO;
HASEGAWA, YOSHIKAZU;SETO, TOSHIO;OSAWA, TOSHIO
TITLE OF INVENTION: RECOMBINANT NATURAL KILLER CELL ACTIVATOR
NUMBER OF SEQUENCES: 31
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/07/392,841
FILING DATE: 11-AUG-1989

```

SEQ ID NO:20:
LENGTH: 9
5185431-20

Query Match 1.8%; Score 6; DB 6; Length 9;
Best Local Similarity 100.0%; Pred. No. 26+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 266 QWVDS 271
111111
DB 3 QWVDS 8

RESULT 26
US-08-480-229C-12
Sequence 12, Application US/08480229C
Patent No. 5874562
GENERAL INFORMATION:
APPLICANT: Quettermous, Thomas
APPLICANT: Hogan, Bridgid
APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pennie & Edmunds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,229C
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0026-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TEXT: 66141 Pennie
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-229C-12
Query Match 1.8%; Score 6; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
QY 271 SPLSPS 276
111111
DB 1 SPLSPS 6

APPLICANT: Snodgrass, H. Ralph
APPLICANT: Zupancic, Thomas J.
TITLE OF INVENTION: DEVELOPMENTALLY-REGULATED ENDOTHELIAL
TITLE OF INVENTION: CELL LOCUS-1
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pennie & Edmunds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,235C
FILING DATE: 05-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462
REFERENCE/DOCKET NUMBER: 8907-0044-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869 8864/9741
TEXT: 66141 Pennie
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-659-235C-12
Query Match 1.8%; Score 6; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0;
QY 271 SPLSPS 276
111111
DB 1 SPLSPS 6

RESULT 28
US-07-908-317-10
Sequence 10, Application US/07908317
Patent No. 5420027
GENERAL INFORMATION:
APPLICANT: FISHER, CHARLES W.
APPLICANT: BARNES, HENRY J.
APPLICANT: ESTABROOK, RONALD W.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
TITLE OF INVENTION: THE EXPRESSION OF BIOLOGICALLY
TITLE OF INVENTION: ACTIVE FUSION PROTEINS COMBINED
TITLE OF INVENTION: EPITOPIC CYTOCHROME P450 FUSION
TITLE OF INVENTION: A REDUCTASE IN BACTERIA
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: ARNOLD, WHITE & BURKE
STREET: P.O. BOX 4433
CITY: HOUSTON
STATE: TEXAS
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/908,317
 FILING DATE: 19920702
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: PARKER, DAVID L.
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: UTSD:292/PAR
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512-320-7200
 TELEFAX: 512-474-7577
 TELETYPE: NOT APPLICABLE
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 19 amino acid residues
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-07-908-317-10

Query Match 1.8%; Score 6; DB 1; Length 19;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 ESTIML 33
 111111
 Db 11 ESTIML 16

RESULT 29
 PCT-US93-06171-10
 Sequence 10, Application PCT/US9306171
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: FUSION PROTEINS COMPRISING
 TITLE OF INVENTION: EUKARYOTIC CYTOCHROME P450 FUSFD TO
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ARNOID, WHITE & DURREE
 STREET: P.O. BOX 4433
 CITY: HOUSTON
 STATE: TEXAS
 COUNTRY: USA
 ZIP: 77210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/06171
 FILING DATE: 19930629
 CLASSIFICATION:
 PRIOR APPLICATION NUMBER:
 APPLICATION DATA:
 FILING DATE: 07/908,317
 ATTORNEY/AGENT INFORMATION:
 NAME: PARKER, DAVID L.
 REGISTRATION NUMBER: 32,165
 REFERENCE/DOCKET NUMBER: UTSD:292/PAR
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 512-320-7200
 TELEFAX: 512-474-7577
 TELETYPE: NOT APPLICABLE
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 19 amino acid residues
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 PCT-US93-06171-10
 Query Match 1.8%; Score 6; DB 5; Length 19;

Best Local Similarity 100.0%; Pred. No. 45;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 ESTIML 33
 111111
 Db 11 ESTIML 16

RESULT 30
 US-08-839-624-24
 Sequence 24, Application US/08839624
 Patent No. 625045
 GENERAL INFORMATION:
 APPLICANT: Kohn et al.
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
 NUMBER OF SEQUENCES: 43
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Banner & Witcoff, Inc.
 STREET: One Financial Center
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02111
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WordPerfect 6.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08839,624
 FILING DATE: April 15, 1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/72151
 FILING DATE: 15-APR-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/017,268
 FILING DATE: 13-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Kathleen M. Williams
 REGISTRATION NUMBER: 34,380
 REFERENCE/DOCKET NUMBER: 4255/5390
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-345-9100
 TELEFAX: 617-345-9111
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 25 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 US-08-839-624-24
 Query Match 1.8%; Score 6; DB 4; Length 25;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 315 ESTIML 321
 111111
 Db 20 KRPAC 25

RESULT 31
 US-09-150-812-24
 Sequence 24, Application US/09150812
 Patent No. 649891
 GENERAL INFORMATION:
 APPLICANT: Kohn et al.
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
 HIV INFECTION
 NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:
ADDRESSER: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,812
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-150-812-24

Query Match 1.8%, Score 6, DB 4, Length 25;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 KKPAC 321
11111
DB 20 KKPAC 25

RESULT 32
US-08-839-624-23
Sequence 23, Application US/08839624
Patent No. 6225045
GENERAL INFORMATION:
APPLICANT: Karn et al.
TITLE OF INVENTION: METHOD AND COMPOSITIONS FOR COMBATING
NUMBER OF SEQUENCES: 43
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,624
FILING DATE: April 15, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: P01/3896/78191
FILING DATE: 15-Apr-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-839-624-23

Query Match 1.8%, Score 6, DB 4, Length 33;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

OY 316 KKPAC 321
11111
DB 28 KKPAC 33

RESULT 33
US-09-150-812-23
Sequence 23, Application US/09150812
Patent No. 6395891
GENERAL INFORMATION:
APPLICANT: Karn et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
NUMBER OF SEQUENCES: 43
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,812
FILING DATE: 11-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: 13-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown

MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-150-812-23

Query Match 1.8%; Score 6; DB 4; Length 33;
Best Local Similarity 100.0%; Prod No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 KRPAC 321
Db 28 KRPAC 33

RESULT 34
US-08-839-624-25
Sequence 25, Application US/08839624
Patent No. 6225045
GENERAL INFORMATION:
APPLICANT: Kain et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
TITLE OF INVENTION: HIV INFECTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/839,624
FILING DATE: April 15, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/78191
FILING DATE: 15-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-839-624-25

Query Match 1.8%; Score 6; DB 4; Length 37;
Best Local Similarity 100.0%; Prod. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 KRPAC 321
Db 32 KRPAC 37

RESULT 35
US-09-150-812-25
Sequence 25, Application US/09150812
Patent No. 6395891

GENERAL INFORMATION:
APPLICANT: Kain et al.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
TITLE OF INVENTION: HIV INFECTION
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Witcoff, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,812
FILING DATE: 11-SEP-1998
CLASSIFICATION: <UNKNOWN>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: <UNKNOWN>
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 37 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-150-812-25

Query Match 1.8%; Score 6; DB 4; Length 37;
Best Local Similarity 100.0%; Prod. No. 85;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 KRPAC 321
Db 32 KRPAC 37

RESULT 36
US-08-189-331-143
Sequence 143, Application US/08189331
Patent No. 5747334
GENERAL INFORMATION:
APPLICANT: Kay, H. K.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 186
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penrite & Edmunds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/189,331
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-155
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 143:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-189-331-143

Query Match 1.8% Score 6: DB 1: Length 39:
Best Local Similarity 100.0% Pred. No. 89:
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 250 RPSMWG 255
DB 6 RPSMWG 11

RESULT 37
US-08-471-068-143
Sequence 143, Application US/08471068
Patent No. 5948635
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowlkes, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 186
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penale & Edmunds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,068
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/189,331
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-155
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 143:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide

US-08-471-068-143

Query Match 1.8% Score 6: DB 2: Length 39:
Best Local Similarity 100.0% Pred. No. 89:
Matches 6: Conservative 0: Mismatches 0: Indels 0:

QY 250 RPSMWG 255
DB 6 RPSMWG 11

RESULT 38
US-08-689-421-1
Sequence 1, Application US/08689421
Patent No. 6008029
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie S.
APPLICANT: Brown, Kimberly M.
APPLICANT: Kauppinen, Sakari
APPLICANT: Halkier, Torben P.
TITLE OF INVENTION: PURIFIED COPOLYMERS LACTASES AND NUC
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 60080290 No. 6008029 disk of No. 6008029
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/689,421
FILING DATE: 9-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Harrington, James J.
REGISTRATION NUMBER: 38,711
REFERENCE/DOCKET NUMBER: 4554,204-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0124
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-689-421-1

Query Match 1.8% Score 6: DB 3: Length 40:
Best Local Similarity 100.0% Pred. No. 91:
Matches 6: Conservative 0: Mismatches 0: Indels 0:

QY 309 VDNWVI 314
DB 33 VDNWVI 38

RESULT 39
US-09-389-528-1
Sequence 1, Application US/09389528
Patent No. 6207430
GENERAL INFORMATION:
APPLICANT: Yaver, Debbie S.
APPLICANT: Brown, Kimberly M.
APPLICANT: Kauppinen, Sakari
APPLICANT: Halkier, Torben P.

```

? TITLE OF INVENTION: PURIFIED COPPINUS LACCASES AND NUCLEIC ACIDS
? TITLE OF INVENTION: ENCODING SAME
? NUMBER OF SEQUENCES: 36
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: No 6207430- No 6207430disk of No 6207430th America, Inc.
? STREET: 405 Lexington Avenue, 64th Floor
? CITY: New York
? STATE: New York
? COUNTRY: United States of America
? ZIP: 10174-6401
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent Release #1 0, Version #1 30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/389,528
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/98/689,421
? FILING DATE: 9-AUG-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Harrington, James J.
? REGISTRATION NUMBER: 38,711
? REFERENCE/DOCKET NUMBER: 4554,204 WO
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 212-867-0123
? TELEFAX: 212-878-9655
? INFORMATION FOR SEQ ID NO 1
? SEQUENCE CHARACTERISTICS:
? LENGTH: 40 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? US-09-389-528-1

Query Match 1.8%; Score 6; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 309 VDNWMI 314
|11111
DB 33 VDNWMI 38

RESULT 40
US-09-181-827A-1
? Sequence 1, Application US/09181827A
? Patent No. 6242232
? GENERAL INFORMATION:
? APPLICANT: Yaver, Debbie S.
? APPLICANT: Brown, Kimberly M.
? APPLICANT: Kaupinen, Sakari
? APPLICANT: Halkier, Torben P.
? TITLE OF INVENTION: Purified Coprinus laccases And Nucleic
? FILE REFERENCE: 4554,200-US
? CURRENT APPLICATION NUMBER: US/09/181,827A
? CURRENT FILING DATE: 1998-10-28
? PRIOR APPLICATION NUMBER: 60/002,800
? PRIOR FILING DATE: 1995-08-25
? NUMBER OF SEQ ID NOS: 36
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 1
? LENGTH: 40
? TYPE: PRT
? ORGANISM: Coprinus cinereus
? US-09-181-827A-1

Query Match 1.8%; Score 6; DB 4; Length 40;
Best Local Similarity 100.0%; Pred No 91;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 309 VDNWMI 314
|11111
DB 33 VDNWMI 38

RESULT 41
US-08-839-624-26
? Sequence 26, Application US/08839624
? Patent No. 625045
? GENERAL INFORMATION:
? APPLICANT: Karn et al.
? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
? NUMBER OF SEQUENCES: 43
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Hammer & Wilcott, Inc.
? STREET: One Financial Center
? CITY: Boston
? STATE: Massachusetts
? COUNTRY: USA
? ZIP: 02111
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Wordperfect 6.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/839,624
? FILING DATE: April 15, 1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/396/78191
? FILING DATE: 15-APR-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/017,268
? FILING DATE: 13-MAY-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Kathleen M. Williams
? REGISTRATION NUMBER: 34,380
? REFERENCE/DOCKET NUMBER: 3255/5390
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-345-9100
? TELEFAX: 617-345-9111
? INFORMATION FOR SEQ ID NO: 25:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 49 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: unknown
? MOLECULE TYPE: peptide
? US-08-839-624-26

Query Match 1.8%; Score 6; DB 4; Length 49;
Best Local Similarity 100.0%; Pred. No. 1,1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 316 KKPAC 421
|11111
DB 44 KKPAC 49

RESULT 42
US-09-150-812-26
? Sequence 26, Application US/09150812
? Patent No. 6305891
? GENERAL INFORMATION:
? APPLICANT: Karn et al.
? TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COMBATING
? HIV INFECTION
? NUMBER OF SEQUENCES: 43
? CORRESPONDENCE ADDRESS:

```



```

ADDRESS: Banner & Wilcott, Inc.
STREET: One Financial Center
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,812
FILING DATE: 11-Sep-1998
CLASSIFICATION: <unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/839,624
FILING DATE: <unknown>
APPLICATION NUMBER: US 60/017,268
FILING DATE: 13-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kathleen M. Williams
REGISTRATION NUMBER: 34,380
REFERENCE/DOCKET NUMBER: 3255/5390
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-345-9100
TELEFAX: 617-345-9111
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-150-812-26

Query Match      1.8% Score 6; DB 4; Length 49;
Best Local Similarity 100.0%; Pred No. 1 1e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 316 KKPAAAC 321
Db 44 KKPAAAC 49

RESULT 44
US-08-936-165A-308
Sequence 308, Application US/08936165A
Patent No. 6348582
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: No. 6348582: Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESS:
ADDRESSEE: Smithkline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette

```

```

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/027,072
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50649
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 408:
SEQUENCE CHARACTERISTICS:
LENGTH: 52 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-936-165A-308

Query Match      1.8% Score 6; DB 4; Length 52
Best Local Similarity 100.0%; Pred No. 1 2e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

OY 224 EVRAQL 229
Db 17 EVRAQL 22

RESULT 44
US-07-664-989B-78
Sequence 78, Application US/07664989B
Patent No. 5223409
GENERAL INFORMATION:
APPLICANT: Ladner, Robert Charles
APPLICANT: Guterman, Sonia Kosow
APPLICANT: Roberts, Bruce Lindsay
APPLICANT: Markland, William
APPLICANT: Ley, Arthur Charles
TITLE OF INVENTION: Directed Evolution of No. 5223409
TITLE OF INVENTION: Binding Proteins
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brody and Netmark
STREET: 419 Seventh Street, N.W.
STREET: Suite 300
CITY: Washington,
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 4.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07664,989B
FILING DATE: 19910301
CLASSIFICATION: 530
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: P-1,785B,204731
FILING DATE: 01-SEP-1989
PRIORITY APPLICATION DATA: 07/487,064
APPLICATION NUMBER: 02-MAR-1990
FILING DATE: 02-MAR-1990

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/240,160
FILING DATE: 02-SEP-1988
ATTORNEY/AGENT INFORMATION:
NAME: Cooper, Iver P.
REGISTRATION NUMBER: 28005
REFERENCE/DOCKET NUMBER: LADNER 7
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-664-9898-78

Query Match 1.8%; Score 6; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 1,3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GCLGNG 21
|||||
DB 37 GCLGNG 42

RESULT 45
US-08-358-160-55
Sequence 55, Application US/08358160
Patent No. 5663143
GENERAL INFORMATION:
APPLICANT: LEY, Arthur C.
APPLICANT: LADNER, Robert C.
APPLICANT: GUTERMAN, Sonia K.
APPLICANT: ROBERTS, Bruce L.
APPLICANT: MARKLAND, William
APPLICANT: KENT, Rachel H.
TITLE OF INVENTION: ENGINEERED HUMAN-DERIVED KUNITZ
NUMBER OF SEQUENCES: 234
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W. Suite 300
City: Washington
STATE: District of Columbia
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/358,160
FILING DATE: 16-DEC-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/133,031
FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,319
FILING DATE: 26-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/664,989
FILING DATE: 01-MAR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/487,063
FILING DATE: 02-MAR-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/240,160
FILING DATE: 02-SEP-1988
ATTORNEY/AGENT INFORMATION:

NAME: Cooper, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: LEY-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 58 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-358-160-55

Query Match 1.8%; Score 6; DB 1; Length 58;
Best Local Similarity 100.0%; Pred. No. 1,3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GCLGNG 21
|||||
DB 37 GCLGNG 42

Search completed: December 7, 2002, 11:18:14
Job time : 22 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: December 7, 2002, 11:15:39, Search time 13 seconds

(without alignments)
406,058 Million cell updates/sec

Title: US-09-831-458a-12

Perfect score: 125

Sequence: 1 MSMSKPRVVOGLHGLGH PCRWNYWYCKRPAACEFDE 325

Scoring table: OHGO

Gapop 60 0, Gapext 60 0

Searched: 103943 seqs, 16242309 residues

Word size: 5

Total number of hits satisfying chosen parameters: 3135

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database: Published Applications-AA*

1: /cgn2_6/prodata/2/pubpaa/us08_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PC1_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PC107_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/PC107_PUBCOMB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	15.4	50	US-09-864-761-40945	Sequence 40945, A
2	19	2.8	31	US-10-125-452-27	Sequence 27, Appl
3	9	2.8	339	US-09-983-531A-26	Sequence 26, Appl
4	9	2.8	394	US-09-983-531A-16	Sequence 16, Appl
5	9	2.8	427	US-09-983-531A-20	Sequence 20, Appl
6	9	2.8	779	US-10-125-470-15	Sequence 15, Appl
7	9	2.8	778	US-10-125-452-16	Sequence 16, Appl
8	9	2.8	918	US-10-020-733-4	Sequence 4, Appl
9	9	2.8	918	US-10-125-470-9	Sequence 9, Appl
10	9	2.8	918	US-10-125-452-9	Sequence 9, Appl
11	9	2.8	926	US-10-020-733-2	Sequence 2, Appl
12	9	2.8	955	US-10-020-733-8	Sequence 8, Appl
13	9	2.8	963	US-10-020-733-6	Sequence 6, Appl
14	8	2.5	472	US-09-864-582-20	Sequence 20, Appl
15	7	2.2	144	US-09-748-473-127	Sequence 127, Appl
16	7	2.2	199	US-09-825-882-8	Sequence 8, Appl
17	7	2.2	367	US-09-815-242-10118	Sequence 10118, A
18	7	2.2	455	US-09-815-242-14020	Sequence 14020, A
19	7	2.2	548	US-09-732-350-4	Sequence 4, Appl

20	7	2.2	720	US-09-919-497-83	Appl
21	7	2.2	769	US-10-072-841-31	Appl
22	6	1.8	29	US-09-982-172-15	Appl
23	6	1.8	29	US-09-982-172-215	Appl
24	6	1.8	31	US-09-864-761-48338	Appl
25	6	1.8	37	US-09-864-761-39821	Appl
26	6	1.8	38	US-09-864-761-48289	Appl
27	6	1.8	44	US-09-864-761-44451	Appl
28	6	1.8	49	US-09-864-761-38145	Appl
29	5	1.8	53	US-09-864-761-40929	Appl
30	6	1.8	52	US-09-919-980-308	Appl
31	6	1.8	54	US-09-764-860-353	Appl
32	6	1.8	57	US-09-764-877-1727	Appl
33	6	1.8	58	US-09-781-988-78	Appl
34	6	1.8	90	US-09-867-550-1392	Appl
35	6	1.8	102	US-10-004-381-17	Appl
36	6	1.8	111	US-09-764-853-786	Appl
37	6	1.8	117	US-09-864-761-45806	Appl
38	6	1.8	119	US-09-764-870-302	Appl
39	6	1.8	131	US-09-815-242-12092	Appl
40	6	1.8	138	US-09-216-393-66	Appl
41	6	1.8	139	US-09-815-242-10702	Appl
42	6	1.8	153	US-09-764-868-1057	Appl
43	6	1.8	154	US-09-825-301-27	Appl
44	6	1.8	154	US-09-604-2878-47	Appl
45	6	1.8	154	US-10-007-805-471	Appl

ALIGNMENTS

RESULT 1
US-09-864-761-40945
Sequence 40945, Application US-09-864-761
Patent No. US-2004/08763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENE-DERIVED SINGLE EXON NUCLEOTIDE SEQUENCES
FILE REFERENCE: Aomica-X-1
CURRENT FILING DATE: 2001-05-24
PRIOR APPLICATION NUMBER: US-60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US-60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US-09/764,868-1057
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: GR-24264-6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US-60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US-60/200,170-666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US-60/200,667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US-60/200,664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US-60/200,669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US-60/200,665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US-60/200,668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US-60/200,663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US-60/200,662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US-60/200,661
PRIOR FILING DATE: 2001-01-30

```

PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 40945
LENGTH: 50
TYPE: PRI
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC008812.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN FETAL MAPROM, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: SWISSPROT Hit: P26717, EVALUE 6.99e-07
OTHER INFORMATION: EST_HUMAN Hit: H00958.1, EVALUE 2.00e-25
US-09-864-761-40945

```

```

Query Match: 15.4%; Score 50; DB 10; Length 50;
Best Local Similarity 100.0%; Pred. No. 2.8e-30;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 184 PTC/HCEKMTFFGCGYMSNSQPMHDSVACQCEPAQLVVIKTAEEQ 238
DB 1 RICHCHCKDKTFFGCGYMSNSQPMHDSVACQCEPAQLVVIKTAEEQ 50

```

```

RESULT 2
US-10-125-452-27
Sequence 27; Application US/20125452
Patent No. US20020173640A1
GENERAL INFORMATION:
APPLICANT: Ruden et al.
TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and Antibodies
FILE REFERENCE: p76bnp2
CURRENT APPLICATION NUMBER: US/10/125,452
PRIORITY FILING DATE: 2002-04-19
PRIORITY APPLICATION NUMBER: 09/955,504
PRIORITY FILING DATE: 2001-09-19
PRIORITY APPLICATION NUMBER: 09/712,907
PRIORITY FILING DATE: 2000-11-16
PRIORITY APPLICATION NUMBER: PCT/US00/14308
PRIORITY FILING DATE: 2000-05-25
PRIORITY APPLICATION NUMBER: 60/178,717
PRIORITY FILING DATE: 2000-01-28
PRIORITY APPLICATION NUMBER: 60/142,930
PRIORITY FILING DATE: 1999-07-09
PRIORITY APPLICATION NUMBER: 60/136,388
PRIORITY FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 27
LENGTH: 41
TYPE: PRI
ORGANISM: Homo sapiens
US-10-125-452-27

```

```

Query Match: 2.8%; Score 9; DB 9; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.083;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 34 AGYVALIV 42
DB 1 AGYVALIV 11

```

```

RESULT 3
US-09-983-531A-26
Sequence 26; Application US/09983531A
Patent No. US-0920147132A1
GENERAL INFORMATION:
APPLICANT: Fujisawa, Atsuko
APPLICANT: Yamakawa, Toru
APPLICANT: Shitakawa, Kanon
APPLICANT: Chitose, Ofii
APPLICANT: Ogawa, Naoki
TITLE OF INVENTION: Meltrins
FILE REFERENCE: 11-22-99, sequence submission
CURRENT APPLICATION NUMBER: US/09/983,531A
PRIORITY FILING DATE: 2001-10-24
PRIORITY APPLICATION NUMBER: JP 8-61756
PRIORITY FILING DATE: 1996-02-24
PRIORITY APPLICATION NUMBER: PCT/JP96/03017
PRIORITY FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 26
LENGTH: 349
TYPE: PRI
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: "Novo"
OTHER INFORMATION: JMI09, (p61 beta 24c), human meltrin beta
US-09-983-531A-26

```

```

Query Match: 2.8%; Score 9; DB 10; Length 339;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 34 AGYVALIV 42
DB 245 AGYVALIV 253

```

```

RESULT 4
US-09-983-531A-16
Sequence 16; Application US/09983531A
Patent No. US20020147132A1
GENERAL INFORMATION:
APPLICANT: Fujisawa, Atsuko
APPLICANT: Yamakawa, Toru
APPLICANT: Shitakawa, Kanon
APPLICANT: Chitose, Ofii
APPLICANT: Ogawa, Naoki
TITLE OF INVENTION: Meltrins
FILE REFERENCE: 11-22-99, sequence submission
CURRENT APPLICATION NUMBER: US/09/983,531A
PRIORITY FILING DATE: 2001-10-24
PRIORITY APPLICATION NUMBER: JP 8-61756
PRIORITY FILING DATE: 1996-02-24
PRIORITY APPLICATION NUMBER: PCT/JP96/03017
PRIORITY FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 16
LENGTH: 394
TYPE: PRI
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Human meltrin
OTHER INFORMATION: beta derived from cDNA
US-09-983-531A-16

```

```

Query Match: 2.8%; Score 9; DB 10; Length 394;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 34 AGVLVAILV 42
|||||
Db 333 AGVLVAILV 341

RESULT 5

US-09-983-531A-20
Sequence 20, Application US/09983531A
Patent No. US20020147132A1
GENERAL INFORMATION:
APPLICANT: Fujisawa, Atsuko
APPLICANT: Yamakawa, Toru
APPLICANT: Shirakawa, Kamon
APPLICANT: Chitose, Oriti
APPLICANT: Ogawa, Naoki
TITLE OF INVENTION: Meltrins
FILE REFERENCE: 11-22-99 sequence submission
CURRENT APPLICATION NUMBER: US/09/983,531A
CURRENT FILING DATE: 2001-10-24
PRIOR APPLICATION NUMBER: JP 8-61756
PRIOR FILING DATE: 1996-02-23
PRIOR APPLICATION NUMBER: PCT/JP96/03017
PRIOR FILING DATE: 1996-10-17
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 20
LENGTH: 427
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Clones:
OTHER INFORMATION: JM109(pmel-beta-24C) and JM109(pmel-beta-24N),
US-09-983-531A-20

Query Match 2.8%; Score 9, DB 10, Length 427;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 AGVLVAILV 42
|||||
Db 333 AGVLVAILV 341

RESULT 6

US-10-125-470-16
Sequence 16, Application US/10125470
Patent No. US20020165377A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and Antibodies
FILE REFERENCE: PTO06P1
CURRENT APPLICATION NUMBER: US/10/125,470
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US/09/712,907A
PRIOR FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: 50/130,588
PRIOR FILING DATE: 1999-05-27
PRIOR APPLICATION NUMBER: 60/142,930
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/178,717
PRIOR FILING DATE: 2000-01-28
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16
LENGTH: 778
TYPE: PRT
ORGANISM: Homo sapiens
US-10-125-470-16

Query Match 2.8%; Score 9, DB 9, Length 778;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 AGVLVAILV 42
|||||
Db 705 AGVLVAILV 713

RESULT 7

US-10-125-452-16
Sequence 16, Application US/10125472
Patent No. US20020173640A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and
FILE REFERENCE: PTO06P2
CURRENT APPLICATION NUMBER: US/10/125,452
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 09/955,504
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: 09/712,607
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: PCT/US00/14408
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/178,717
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/142,930
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/136,088
PRIOR FILING DATE: 1999-05-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 16
LENGTH: 778
TYPE: PRT
ORGANISM: Homo sapiens
US-10-125-452-16

Query Match 2.8%; Score 9, DB 9, Length 778;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0;

OY 34 AGVLVAILV 42
|||||
Db 705 AGVLVAILV 713

RESULT 8

US-10-020-733-4
Sequence 4, Application US/10020733
Patent No. US2002015214A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wilgowski, Nathaniel L.
TITLE OF INVENTION: Fridtle, Carl Johan
FILE REFERENCE: LEX-0263-058A
CURRENT APPLICATION NUMBER: US/10/020,733
CURRENT FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 09/60/244,999
PRIOR FILING DATE: 2000-11-01
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 918
TYPE: PRT
ORGANISM: homo sapiens
US-10-020-733-4

Query Match 2.8%; Score 9, DB 9, Length 918;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0;

OY 34 AGVLVAILV 42
|||||

DB 705 ACVLAVALIV 713

```

RESULT 9
US-10-125-470-9
: Sequence 9, Application US/10125470
: Patent No. US2002016377A1
: GENERAL INFORMATION:
: APPLICANT: Ruben et al.
: TITLE OF INVENTION: ADAM Polymorphoelides, Polypeptides, and Antibodies
: FILE REFERENCE: PT006P1
: CURRENT APPLICATION NUMBER: US/10/125,470
: PRIOR FILING DATE: 2000-04-19
: PRIOR APPLICATION NUMBER: US/09/712,907A
: PRIOR FILING DATE: 2001-03-07
: PRIOR APPLICATION NUMBER: 60/136,388
: PRIOR FILING DATE: 1999-06-27
: PRIOR APPLICATION NUMBER: 60/142,930
: PRIOR FILING DATE: 1999-07-09
: PRIOR APPLICATION NUMBER: 60/178,717
: PRIOR FILING DATE: 2000-01-28
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9
: LENGTH: 918
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-125-470-9

```

Query Match 2.8%; Score 9; DB 9; Length 918;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 ACVLAVALIV 42
 DB 705 ACVLAVALIV 713

```

RESULT 10
US-10-125-452-9
: Sequence 9, Application US/10125452
: Patent No. US20020173640A1
: GENERAL INFORMATION:
: APPLICANT: Ruben et al.
: TITLE OF INVENTION: ADAM Polymorphoelides, Polypeptides, and Antibodies
: FILE REFERENCE: PT006P2
: CURRENT APPLICATION NUMBER: US/10/125,452
: PRIOR FILING DATE: 2002-04-19
: PRIOR APPLICATION NUMBER: 09/055,504
: PRIOR FILING DATE: 2001-09-19
: PRIOR APPLICATION NUMBER: 09/712,907
: PRIOR FILING DATE: 2000-11-16
: PRIOR APPLICATION NUMBER: PCT/US00/14108
: PRIOR FILING DATE: 2000-05-25
: PRIOR APPLICATION NUMBER: 60/178,717
: PRIOR FILING DATE: 2000-01-28
: PRIOR APPLICATION NUMBER: 60/142,930
: PRIOR FILING DATE: 1999-07-09
: PRIOR APPLICATION NUMBER: 60/136,388
: PRIOR FILING DATE: 1999-06-27
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9
: LENGTH: 918
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-125-452-9

```

Query Match 2.8%; Score 9; DB 9; Length 918;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 ACVLAVALIV 42

DB 705 ACVLAVALIV 713

```

RESULT 11
US-10-020-733-2
: Sequence 2, Application US/10020733
: Patent No. US20020161214A1
: GENERAL INFORMATION:
: APPLICANT: Witke, D. Wade
: APPLICANT: Wilgowski, Nathaniel L.
: TITLE OF INVENTION: No. US20020161214A1 Human Proteases and Polymorphoelides Enco
: FILE REFERENCE: LEX-0263-USA
: CURRENT APPLICATION NUMBER: US/10/020,733
: PRIOR FILING DATE: 2001-10-10
: PRIOR APPLICATION NUMBER: US 60/244,939
: PRIOR FILING DATE: 2000-11-01
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 926
: TYPE: PRT
: ORGANISM: homo sapiens
US-10-020-733-2

```

Query Match 2.8%; Score 5; DB 9; Length 926;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 ACVLAVALIV 42
 DB 705 ACVLAVALIV 713

```

RESULT 12
US-10-020-733-8
: Sequence 8, Application US/10020733
: Patent No. US20020161214A1
: GENERAL INFORMATION:
: APPLICANT: Witke, D. Wade
: APPLICANT: Wilgowski, Nathaniel L.
: APPLICANT: Fridley, Carl Johan
: TITLE OF INVENTION: No. US20020161214A1 Human Proteases and Polymorphoelides Enco
: FILE REFERENCE: LEX-0263-USA
: CURRENT APPLICATION NUMBER: US/10/020,733
: PRIOR FILING DATE: 2001-10-10
: PRIOR APPLICATION NUMBER: US 60/244,939
: PRIOR FILING DATE: 2000-11-01
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 8
: LENGTH: 955
: TYPE: PRT
: ORGANISM: homo sapiens
US-10-020-733-8

```

Query Match 2.8%; Score 9; DB 9; Length 955;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 34 ACVLAVALIV 42
 DB 705 ACVLAVALIV 713

```

RESULT 13
US-10-020-733-6
: Sequence 6, Application US/10020733
: Patent No. US20020161214A1
: GENERAL INFORMATION:
: APPLICANT: Witke, D. Wade
: APPLICANT: Wilgowski, Nathaniel L.

```

```

: APPLICANT: Fiddle, Carl Johan
: TITLE OF INVENTION: NO. US20020161214A1e1 Human Proteases and Peptidomimetics Encoodin
: FILE REFERENCE: LEX-0263-085A
: CURRENT APPLICATION NUMBER: US/10/020,733
: CURRENT FILING DATE: 2001-10-30
: PRIOR APPLICATION NUMBER: US 60/244,939
: PRIOR FILING DATE: 2000-11-01
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6
: LENGTH: 963
: TYPE: PRT
: ORGANISM: homo sapiens
: US-10-020-733-6

Query Match
Best Local Similarity 100.0%; Score 9; DB 9; Length 963;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AGVLAIV 42
DB 705 AGVLAIV 713

RESULT 14
US-09-866-582-29
: Sequence 29, Application US/09866582
: Patent No. US20020127620A1
: GENERAL INFORMATION:
: APPLICANT: Witman, George R.
: APPLICANT: Pacout, Gregory J.
: APPLICANT: Rosenbaum, Joel L.
: APPLICANT: Cole, Douglas G.
: TITLE OF INVENTION: INTRAFACELLAR TRANSPORT
: FILE REFERENCE: 07917-145001
: CURRENT APPLICATION NUMBER: US/09/866,582
: CURRENT FILING DATE: 2001-05-24
: PRIOR APPLICATION NUMBER: US 60/206,923
: PRIOR FILING DATE: 2000-05-24
: NUMBER OF SEQ ID NOS: 50
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 29
: LENGTH: 472
: TYPE: PRT
: ORGANISM: Caenorhabditis elegans
: US-09-866-582-29

Query Match
Best Local Similarity 100.0%; Score 8; DB 10; Length 472;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 DQSKQOQI 176
DB 21 DQSKQOQI 28

RESULT 15
US-09-738-973-427
: Sequence 427, Application US/09738973
: Patent No. US20020110563A1
: GENERAL INFORMATION:
: APPLICANT: Reed, Steven G.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Lodes, Michael J.
: APPLICANT: Fling, Steven P.
: APPLICANT: Mohamath, Radoch
: APPLICANT: Algate, Paul A.
: APPLICANT: Secrist, Heather
: APPLICANT: Indrias, Carol Joseph
: APPLICANT: Renson, Darin K.
: APPLICANT: Elliot, Mark
: APPLICANT: Mannion, Jane
: APPLICANT: Kalos, Michael D.
```

```

: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
: TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CAN
: FILE REFERENCE: 210121.47569
: CURRENT APPLICATION NUMBER: US/09/738,973
: CURRENT FILING DATE: 2000-12-14
: NUMBER OF SEQ ID NOS: 587
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 427
: LENGTH: 184
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-738-973-427

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 18
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 QEVRAQL 229
DB 162 QEVRAQL 168

RESULT 16
US-09-825-882-8
: Sequence 8, Application US/09825882
: Patent No. US20020094551A1
: GENERAL INFORMATION:
: APPLICANT: Adler, Jon Elliot
: TITLE OF INVENTION: T2R TASTE RECEPTORS AND GENES ENCODING
: FILE REFERENCE: 078063/0274527EXT
: CURRENT APPLICATION NUMBER: US/09/825,882
: CURRENT FILING DATE: 2001-04-05
: PRIOR APPLICATION NUMBER: 60/195,532
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: 60/247,014
: PRIOR FILING DATE: 2000-11-13
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: PatchIn Ver. 2.1
: SEQ ID NO 8
: LENGTH: 109
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-825-882-8

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 1
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 LLSFML 33
DB 192 LLSFML 198

RESULT 17
US-09-815-242-10118
: Sequence 10118, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Frawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert L.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
```

```

1 PRIOR APPLICATION NUMBER: 60/206,848
2 PRIOR FILING DATE: 2000-05-23
3 PRIOR APPLICATION NUMBER: 60/207,727
4 PRIOR FILING DATE: 2000-05-26
5 PRIOR APPLICATION NUMBER: 60/242,578
6 PRIOR FILING DATE: 2000-10-23
7 PRIOR APPLICATION NUMBER: 60/253,625
8 PRIOR FILING DATE: 2000-11-27
9 PRIOR APPLICATION NUMBER: 60/257,931
10 PRIOR FILING DATE: 2000-12-22
11 PRIOR APPLICATION NUMBER: 60/269,308
12 PRIOR FILING DATE: 2001-02-16
13 NUMBER OF SEQ ID NOS: 14110
14 SOFTWARE: FastSeq for Windows Version 4.0
15 SEQ ID NO 10118
16 LENGTH: 467
17 TYPE: PRT
18 ORGANISM: Escherichia coli
19 US-09-815-242-10118

```

```

Query Match
Best Local Similarity 2.2%, Score 7, DB 10, Length 367,
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

CY 23 LVGLLS 29
DB 95 LVGLLS 101

```

```

RESULT 18
US-09-815-242-14020
1 Sequence 14020, Application US/00815442
2 Patent No. US20020061569A1
3 GENERAL INFORMATION:
4 APPLICANT: Haselbeck, Robert
5 APPLICANT: Ohlson, Karl L.
6 APPLICANT: Zyskind, Judith W
7 APPLICANT: Wall, Daniel
8 APPLICANT: Trawick, John D
9 APPLICANT: Carr, Grant J.
10 APPLICANT: Yamamoto, Robert T.
11 APPLICANT: Xu, H. Howard
12 TITLE OF INVENTION: Identification of Essential Genes in
13 FILE REFERENCE: ELITRA, 011A
14 CURRENT APPLICATION NUMBER: US/00815442
15 CURRENT FILING DATE: 2001-03-21
16 PRIOR APPLICATION NUMBER: 60/191,078
17 PRIOR FILING DATE: 2000-03-21
18 PRIOR APPLICATION NUMBER: 60/206,848
19 PRIOR FILING DATE: 2000-05-23
20 PRIOR APPLICATION NUMBER: 60/207,727
21 PRIOR FILING DATE: 2000-05-26
22 PRIOR APPLICATION NUMBER: 60/242,578
23 PRIOR FILING DATE: 2000-10-23
24 PRIOR APPLICATION NUMBER: 60/253,625
25 PRIOR FILING DATE: 2000-11-27
26 PRIOR APPLICATION NUMBER: 60/257,931
27 PRIOR FILING DATE: 2000-12-22
28 PRIOR APPLICATION NUMBER: 60/269,308
29 PRIOR FILING DATE: 2001-02-16
30 NUMBER OF SEQ ID NOS: 14110
31 SOFTWARE: FastSeq for Windows Version 4.0
32 SEQ ID NO 14020
33 LENGTH: 455
34 TYPE: PRT
35 ORGANISM: Salmonella typhi
36 US-09-815-242-14020

```

```

Query Match
Best Local Similarity 2.2%, Score 7, DB 10, Length 455,
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

CY 23 LVGLLS 29
DB 183 LVGLLS 189

```

```

RESULT 19
US-09-732-350-4
1 Sequence 4, Application US/00732350
2 Patent No. US20010041490A1
3 GENERAL INFORMATION:
4 APPLICANT: Sorenson, Allan
5 APPLICANT: Xu, Feng
6 TITLE OF INVENTION: LACCASE MUTANTS
7 NUMBER OF SEQUENCES: 10
8 CORRESPONDING ADDRESS:
9 ADDRESS: No. US20010041490A1; No. US20010041490A1; No. US20010041490A
10 SHEET: 405 Lexington Avenue
11 CITY: New York
12 STATE: NY
13 COUNTRY: USA
14 ZIP: 10174
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Diskette
17 COMPUTER: IBM Compatible
18 OPERATING SYSTEM: DOS
19 SOFTWARE: FastSeq for Windows Version 2.0
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/00732,350
22 FILING DATE:
23 CLASSIFICATION:
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: 67/642,315
26 FILING DATE:
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Kozek, Carol
29 REGISTRATION NUMBER: 36,993
30 REFERENCE/INVENT NUMBER: 6200, 200-05
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: 212-867-0123
33 TELEFAX: 212-878-9655
34 INFORMATION FOR SEQ ID NO: 4:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 548 amino acids
37 TYPE: amino acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: protein
41 US-09-732-350-4

```

```

Query Match
Best Local Similarity 2.2%, Score 7, DB 10, Length 548;
Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

CY 308 DIVINYL 314
DB 275 DIVINYL 281

```

```

RESULT 20
US-09-919-497-83
1 Sequence 83, Application US/00919497
2 Patent No. US2002010662A1
3 GENERAL INFORMATION:
4 APPLICANT: Muller, George L.
5 TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
6 FILE REFERENCE: B9801/7225
7 CURRENT APPLICATION NUMBER: US/009194,497
8 CURRENT FILING DATE: 2001-07-31
9 PRIOR APPLICATION NUMBER: US/00721,735
10 PRIOR FILING DATE: 2000-07-31
11 NUMBER OF SEQ ID NOS: 100
12 SOFTWARE: Patcom in version 3.0
13 SEQ ID NO 84

```


LENGTH: 720
TYPE: PRT
ORGANISM: Homo sapiens
US-09-919-497-83

Query Match 2.28: Score 7: PR 10: Length 720:
Best Local Similarity 100.0%: Pred. No. 88:
Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 269 DCSPLSP 275
|||||
DB 247 DCSPLSP 243

RESULT 21
US-10-072-841-31
Sequence 31, Application US/10072841
Patent No. US20020164708A1
GENERAL INFORMATION:

APPLICANT: Sheppard, Dean
Quaranta, Vito
Pytel, Robert
TITLE OF INVENTION: A No. US20020164708A1: Integrin Beta Subunit and Uses Thereof
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:

ADDRESSEE: Preddy, Schroeder, Brueggemann & Clark
STREET: 4470 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92132

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1 0, Version #1 25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/072,841
FILING DATE: 06-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/728,215
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P31 8717
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 769 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-10-072-841-31

Query Match 2.28: Score 7: DB 9: Length 769:
Best Local Similarity 100.0%: Pred. No. 93:
Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 71 AVGEISE 77
|||||
DB 341 AVGEISE 347

RESULT 22
US-09-982-172-16
Sequence 16, Application US/09982172
Patent No. US20020137119A1

GENERAL INFORMATION:
APPLICANT: Emil Israel Katz
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES
TITLE OF INVENTION: DIRECTED THERAPY AGAINST, AND METHODS,
TITLE OF INVENTION: UTILIZING EACH
FILE REFERENCE: 01/22283
CURRENT APPLICATION NUMBER: 95/097042, 172
CURRENT FILING DATE: 2001-10-19
NUMBER OF SEQ ID NOS: 253
SOFTWARE: Patent In version 4.1
SEQ ID NO: 16
LENGTH: 29
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-16

Query Match 1.88: Score 6: DB 10: Length 29:
Best Local Similarity 100.0%: Pred. No. 44:
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 34 AGVLA 39
|||||
DB 10 AGVLA 15

RESULT 23
US-09-982-172-215
Sequence 215, Application US/09982172
Patent No. US20020137119A1
GENERAL INFORMATION:
APPLICANT: Emil Israel Katz
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES
TITLE OF INVENTION: DIRECTED THERAPY AGAINST, AND METHODS,
TITLE OF INVENTION: UTILIZING EACH
FILE REFERENCE: 01/22283
CURRENT APPLICATION NUMBER: 95/097042, 172
CURRENT FILING DATE: 2001-10-19
NUMBER OF SEQ ID NOS: 253
SOFTWARE: Patent In version 4.1
SEQ ID NO: 215
LENGTH: 29
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-215

Query Match 1.88: Score 6: DB 10: Length 29:
Best Local Similarity 100.0%: Pred. No. 44:
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 34 AGVLA 39
|||||
DB 10 AGVLA 15

RESULT 24
US-09-864-761-48348
Sequence 48348, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEOTIDE SEQUENCES
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAYS
FILE REFERENCE: Acomica-X-1
CURRENT APPLICATION NUMBER: US/09864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US-09-864,761

```

PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/287,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/642,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GR 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 48338
LENGTH: 31
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL121923.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
OTHER INFORMATION: EST_HUMAN HTT: A1298543.1, EVALUO 4.006 01
US-09-864-761-48338

```

```

Query Match          1.8%; Score 6; DB 10; Length 31;
Best Local Similarity 100.0%; Pred No 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 241 VIKTAE 236
|||||
DB 15 VIKTAE 20

```

```

RESULT 25
US-09-864-761-39821
Sequence 39821, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Acemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456

```

```

PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/642,356
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GR 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 39821
LENGTH: 37
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004125.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.9
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.9
US-09-864-761-39821

```

```

Query Match          1.8%; Score 6; DB 10; Length 37;
Best Local Similarity 100.0%; Pred No 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 12 LGIGAT 17
|||||
DB 24 LGIGAT 29

```

```

RESULT 26
US-09-864-761-48289
Sequence 48289, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Acemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761

```

```

: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 60/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/246,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 03/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 48289
: LENGTH: 38
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC019248.3
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
: OTHER INFORMATION: SWISSPROT HIT: Q15404, EVALUE 2.00e-15
: OTHER INFORMATION: EST_HUMAN HIT: BF65928.1, EVALUE 3.00e-14
US-09-864-761-48289
Query Match 1.8%, Score 6, DB 10, Length 38;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 87 ERTOLK 92
DB 17 ERTOLK 22
RESULT 27
US-09-864-761-44461
: Sequence 44461, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aeonica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
```

```

: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 60/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/246,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 03/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 44461
: LENGTH: 44
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC06477.2
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.94
: OTHER INFORMATION: EST_HUMAN HIT: AA179279.1, EVALUE 3.1
: OTHER INFORMATION: SWISSPROT HIT: P38613, EVALUE 2.80e-01
US-09-864-761-44461
Query Match 1.8%, Score 6, DB 10, Length 44;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 KVPSSL 51
DB 5 KVPSSL 10
RESULT 28
US-09-864-761-38145
: Sequence 38145, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUC
```

```

1  TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
2  FILE REFERENCE: Acomica-X-1
3  CURRENT FILING DATE: 2001-05-23
4  PRIOR APPLICATION NUMBER: US 60/180,312
5  PRIOR FILING DATE: 2000-02-04
6  PRIOR APPLICATION NUMBER: US 60/207,456
7  PRIOR FILING DATE: 2000-05-26
8  PRIOR APPLICATION NUMBER: US 09/532,366
9  PRIOR FILING DATE: 2000-08-03
10 PRIOR APPLICATION NUMBER: US 24263,6
11 PRIOR FILING DATE: 2000-10-04
12 PRIOR APPLICATION NUMBER: US 60/236,359
13 PRIOR FILING DATE: 2000-09-27
14 PRIOR APPLICATION NUMBER: PCT/US01/00666
15 PRIOR FILING DATE: 2001-01-30
16 PRIOR APPLICATION NUMBER: PCT/US01/00667
17 PRIOR FILING DATE: 2001-01-30
18 PRIOR APPLICATION NUMBER: PCT/US01/00664
19 PRIOR FILING DATE: 2001-01-30
20 PRIOR APPLICATION NUMBER: PCT/US01/00669
21 PRIOR FILING DATE: 2001-01-30
22 PRIOR APPLICATION NUMBER: PCT/US01/00665
23 PRIOR FILING DATE: 2001-01-30
24 PRIOR APPLICATION NUMBER: PCT/US01/00668
25 PRIOR FILING DATE: 2001-01-30
26 PRIOR APPLICATION NUMBER: PCT/US01/00663
27 PRIOR FILING DATE: 2001-01-30
28 PRIOR APPLICATION NUMBER: PCT/US01/00662
29 PRIOR FILING DATE: 2001-01-30
30 PRIOR APPLICATION NUMBER: PCT/US01/00661
31 PRIOR FILING DATE: 2001-01-30
32 PRIOR APPLICATION NUMBER: PCT/US01/00660
33 PRIOR FILING DATE: 2001-01-30
34 PRIOR APPLICATION NUMBER: US 60/234,687
35 PRIOR FILING DATE: 2000-09-21
36 PRIOR APPLICATION NUMBER: US 09/608,408
37 PRIOR FILING DATE: 2000-06-30
38 PRIOR APPLICATION NUMBER: US 09/774,203
39 PRIOR FILING DATE: 2001-01-29
40 NUMBER OF SEQ ID NOS: 49117
41 SOFTWARE: Anomax Sequence Listing Engine v0rs 1.1
42 SEQ ID NO: 36145
43 LENGTH: 49
44 TYPE: PRT
45 ORGANISM: Homo sapiens
46 FEATURE:
47 OTHER INFORMATION: MAP TO A1031026.1
48 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.1
49 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.5
50 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.8
51 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6
52 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.1
53 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.2
54 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4.1
55 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
56 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.9
57 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
58 US-09-864-761-38145
59
60 Query Match: 1.88; Score 6; DB 10; Length 49;
61 Best Local Similarity 100.0%; Pred No 70;
62 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
63
64 QY 271 SPLSPS 276
65 |11111|
66 DB 38 SPLSPS 43
67
68 RESULT 29
69 US-09-864-761-40929
70 : Sequence 40929, Application US/09864761
71 : Patent No. US20020048763A1

```

```

1  GENERAL INFORMATION:
2  APPLICANT: Pharm, Shattou G.
3  APPLICANT: Hankel, David K.
4  APPLICANT: Hankel, David K.
5  APPLICANT: Hankel, David K.
6  TITLE OF INVENTION: HUMAN GENEOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
7  TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
8  FILE REFERENCE: Acomica-X-1
9  CURRENT FILING DATE: 2001-05-23
10 PRIOR APPLICATION NUMBER: US 60/180,312
11 PRIOR FILING DATE: 2000-02-04
12 PRIOR APPLICATION NUMBER: US 60/207,456
13 PRIOR FILING DATE: 2000-05-26
14 PRIOR APPLICATION NUMBER: US 09/532,366
15 PRIOR FILING DATE: 2000-08-03
16 PRIOR APPLICATION NUMBER: US 24263,6
17 PRIOR FILING DATE: 2000-10-04
18 PRIOR APPLICATION NUMBER: US 60/236,359
19 PRIOR FILING DATE: 2000-09-27
20 PRIOR APPLICATION NUMBER: PCT/US01/00666
21 PRIOR FILING DATE: 2001-01-30
22 PRIOR APPLICATION NUMBER: PCT/US01/00667
23 PRIOR FILING DATE: 2001-01-30
24 PRIOR APPLICATION NUMBER: PCT/US01/00664
25 PRIOR FILING DATE: 2001-01-30
26 PRIOR APPLICATION NUMBER: PCT/US01/00669
27 PRIOR FILING DATE: 2001-01-30
28 PRIOR APPLICATION NUMBER: PCT/US01/00665
29 PRIOR FILING DATE: 2001-01-30
30 PRIOR APPLICATION NUMBER: PCT/US01/00668
31 PRIOR FILING DATE: 2001-01-30
32 PRIOR APPLICATION NUMBER: PCT/US01/00663
33 PRIOR FILING DATE: 2001-01-30
34 PRIOR APPLICATION NUMBER: PCT/US01/00662
35 PRIOR FILING DATE: 2001-01-30
36 PRIOR APPLICATION NUMBER: PCT/US01/00661
37 PRIOR FILING DATE: 2001-01-30
38 PRIOR APPLICATION NUMBER: PCT/US01/00660
39 PRIOR FILING DATE: 2001-01-30
40 PRIOR APPLICATION NUMBER: US 60/234,687
41 PRIOR FILING DATE: 2000-09-21
42 PRIOR APPLICATION NUMBER: US 09/608,408
43 PRIOR FILING DATE: 2000-06-30
44 PRIOR APPLICATION NUMBER: US 09/774,203
45 PRIOR FILING DATE: 2001-01-29
46 NUMBER OF SEQ ID NOS: 49117
47 SOFTWARE: Anomax Sequence Listing Engine v0rs 1.1
48 SEQ ID NO: 40929
49 LENGTH: 51
50 TYPE: PRT
51 ORGANISM: Homo sapiens
52 FEATURE:
53 OTHER INFORMATION: MAP TO A1121772.12
54 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
55 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
56 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
57 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
58 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
59 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
60 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
61 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
62 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
63 OTHER INFORMATION: EST_HUMAN HIT: AA360017.1, EVALUATE 4.00e-24
64 US-09-864-761-40929
65
66 Query Match: 1.88; Score 6; DB 10; Length 51;
67 Best Local Similarity 100.0%; Pred No. 73;
68 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
69
70 QY 51 ISOFOS 56
71 |11111|
72 DB 44 ISOFOS 49

```

```

RESULT 30
US-09-939-980-108
: Sequence 308, Application US/09939980
: Patent No. US2002008234A1
GENERAL INFORMATION:
: APPLICANT: Black, Michael
:      Burham, Martin
:      Hodgson, John
:      Knowles, David
:      Lonetto, Michael
:      Nicholas, Richard
:      Pratt, Julie
:      Reichard, Richard
:      Rosenberg, Martin
:      Ward, Judith
: TITLE OF INVENTION: No. US2002008234A1 Polypeptides and Their Uses
:      Polypeptides and Their Uses
: NUMBER OF SEQUENCES: 534
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Smithkline Beecham Corporation
: STREET: 709 Swedeland Road
: CITY: King of Prussia
: STATE: PA
: COUNTRY: USA
: ZIP: 19406-0939
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTING for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/939,980
: FILING DATE: 27-Aug-2001
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/936,165
: FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
: NAME: Gimmil, Edward R
: REGISTRATION NUMBER: 38,891
: REFERENCE/DOCKET NUMBER: P50549
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-270-4478
: TELEFAX: 610-270-5090
: TELEX: <Unknown>
: INFORMATION FOR SEQ ID NO: 308:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 52 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: Protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 308:
US-09-939-980-308

Query Match          1.8%; Score 6; DB 10; Length 52;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 EVRAQL 229
DB 17 EVRAQL 22

RESULT 31
US-09-764-860-353
: Sequence 353, Application US/09764860
: Patent No. US20020094953A1
GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC008

```

```

: CURRENT APPLICATION NUMBER: US/09/764,860
: CURRENT FILING DATE: 2001-01-17
: PRIOR APPLICATION DATA REMOVED
: NUMBER OF SEQ ID NOS: 1198
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 353
: LENGTH: 54
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (12)
: OTHER INFORMATION: Xaa equals any of the naturally occur
: NAME/KEY: SITE
: LOCATION: (47)
: OTHER INFORMATION: Xaa equals any of the naturally occur
: NAME/KEY: SITE
: LOCATION: (52)
: OTHER INFORMATION: Xaa equals any of the naturally occur
US-09-764-860-353

Query Match          1.8%; Score 6; DB 10; Length 54
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 LAGVIV 38
DB 38 LAGVIV 43

RESULT 32
US-09-764-877-1727
: Sequence 1727, Application US/09764877
: Patent No. US20020147140A1
GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC005
: CURRENT APPLICATION NUMBER: US/09/764,877
: CURRENT FILING DATE: 2001-01-17
: PRIOR APPLICATION DATA REMOVED
: NUMBER OF SEQ ID NOS: 4031
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 1727
: LENGTH: 57
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (1)
: OTHER INFORMATION: Xaa equals any of the naturally occur
: NAME/KEY: SITE
: LOCATION: (7)
: OTHER INFORMATION: Xaa equals any of the naturally occur
: NAME/KEY: SITE
: LOCATION: (36)
: OTHER INFORMATION: Xaa equals any of the naturally occur
US-09-764-877-1727

Query Match          1.8%; Score 6; DB 10; Length 57
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QLGILG 16
DB 8 QLGILG 13

RESULT 33
US-09-781-988-78
: Sequence 78, Application US/09781988
: Patent No. US20020150881A1
GENERAL INFORMATION:

```

```

: APPLICANT: Laddner, Robert Charles
: Guterman, Sonia Kosow
: Powers, Prince Lindsay
: Markland, William
: Ley, Arthur Charles
: Kent, Rachel Barbault
: TITLE OF INVENTION: Directed Evolution of No. US2002015081A1a1
: NUMBER OF SEQUENCES: 121
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Browdy and Neimark
: STREET: 419 Seventh Street, N.W.
: Suite 300
: City: Washington,
: STATE: DC
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: WORDPERFECT 4.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/781,988
: FILING DATE: 14-Feb-2001
: CLASSIFICATION: ~Gskcwm~
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 07/664,989
: FILING DATE: ~Gskcwm~
: APPLICATION NUMBER: 07/487,063
: FILING DATE: 02-MAR-1990
: APPLICATION NUMBER: 07/240,160
: FILING DATE: 02-SEP-1988
: ATTORNEY/AGENT INFORMATION:
: NAME: Casper, Iveta P.
: REGISTRATION NUMBER: 28005
: REFERENCE/PROCKET NUMBER: IAINP 7
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202 628 5197
: TELEFAX: 202-737-3528
: INFORMATION FOR SEQ ID NO: 78:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 58 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 78:
US-09-781-988-78

Query Match 1.8% Score 6; PR 10; Length 58;
Best Local Similarity 100.0%; Freq No. 81;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 QCLGCG 21
|111111
DB 37 QCLGCG 42

RESULT 34
US-09-867-550-1392
: Sequence 1392, Application US/9987550
: Patent No. US20020082206A1
: GENERAL INFORMATION:
: APPLICANT: Leach, Martin D.
: APPLICANT: Mehraban, Fuad,
: APPLICANT: Goolley, Pamela
: APPLICANT: Law, Debbie
: APPLICANT: Topper, James
: TITLE OF INVENTION: No. US20020082206A1 Polymorphic peptides from heterogeneous cells and
: FILE REFERENCE: 21402-013 (Gura-313)
: CURRENT APPLICATION NUMBER: US/09/867,550
: CURRENT FILING DATE: 2001-09-20

```

```

: PRIORITY APPLICATION NUMBER: USSN 60/208,427
: PRIORITY FILING DATE: 2000-05-30
: NUMBER OF SEQ ID NOS: 2125
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1392
: LENGTH: 90
: TYPE: prt
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)
: OTHER INFORMATION: N. Merzin Xia may be any one of Ala or Pro as set or Thr
US-09-867-550-1392

Query Match 1.8% Score 6; DB 10; Length 90;
Best Local Similarity 100.0%; Freq No. 1,392+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 SPSTG 276
|111111
DB 81 SPSTG 86

RESULT 35
US-10-004-381-17
: Sequence 17, Application US/10004381
: Patent No. US20020155578A1
: GENERAL INFORMATION:
: APPLICANT: SESTAK, JACK W.
: APPLICANT: WILSON, DAVID S.
: APPLICANT: KEEPE, ANTHONY D.
: TITLE OF INVENTION: STEPTAVIUM-BINDING PEPTIDES AND USES
: FILE REFERENCE: 00786/388002
: CURRENT APPLICATION NUMBER: US/10/004,381
: CURRENT FILING DATE: 2001-10-31
: PRIORITY APPLICATION NUMBER: US 69/244,541
: PRIORITY FILING DATE: 2000-10-31
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 17
: LENGTH: 102
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: selected peptide
US-10-004-381-17

Query Match 1.8% Score 6; DB 9; Length 102;
Best Local Similarity 100.0%; Freq No. 1,38+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 QCLGCG 16
|111111
DB 80 QCLGCG 85

RESULT 36
US-09-764-853-786
: Sequence 786, Application US/09/764854
: Patent No. US20020090672A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, proteins, and Antibodies
: FILE REFERENCE: P1206
: CURRENT APPLICATION NUMBER: US/09/764,853
: CURRENT FILING DATE: 2001-01-17
: PRIORITY APPLICATION data removed - consult PAM or file wrapper
: NUMBER OF SEQ ID NOS: 939
: SOFTWARE: Paton In Ver. 2.0
: SEQ ID NO 786
: LENGTH: 111
: TYPE: PRT

```



```

: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,993
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,309
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12092
: LENGTH: 131
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12092

```

```

Query Match          1.8%; Score 6; DB 10; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 SKEPRV 9
    11111
DB 46 SKEPRV 51

```

```

RESULT 40
US-09-216-393-66
: Sequence 66, Application US/09216393
: Patent No. US2001001447A1
: GENERAL INFORMATION:
: APPLICANT: Milhausen, Michael James
: TITLE OF INVENTION: TOXOPLASMA GONDI1 PROTEINS, NUCLEIC ACID MOLECULES, AND
: FILE REFERENCE: US/09/216,393
: CURRENT APPLICATION NUMBER: US/09/216,393
: PRIOR FILING DATE: 1998-12-18
: EARLIER APPLICATION NUMBER: 08/994,825
: PRIOR FILING DATE: 1997-12-19
: NUMBER OF SEQ ID NOS: 364
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 66
: LENGTH: 138
: TYPE: PRT
: ORGANISM: Toxoplasma gondii
US-09-216-393-66

```

```

Query Match          1.8%; Score 6; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 49 SLSOE 54
    111111
DB 45 SLSOE 50

```

```

RESULT 41
US-09-815-242-10702
: Sequence 10702, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlson, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: PRIOR FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,074
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848

```

```

: PRIOR FILING DATE: 2000-05-24
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-24
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,941
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10702
: LENGTH: 139
: TYPE: PRT
: ORGANISM: Enterococcus faecalis
US-09-815-242-10702

```

```

Query Match          1.8%; Score 6; DB 10; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 86 QELTOL 91
    11111
DB 48 QELTOL 53

```

```

RESULT 42
US-09-764-868-1057
: Sequence 1057, Application US/09764868
: Patent No. US2002016871A1
: GENERAL INFORMATION:
: APPLICANT: Hesse et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Anticodons
: FILE REFERENCE: PTE32
: CURRENT APPLICATION NUMBER: US/09/764,868
: PRIOR FILING DATE: 2001-01-17
: PRIOR APPLICATION data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 1510
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1057
: LENGTH: 153
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-764-868-1057

```

```

Query Match          1.8%; Score 6; DB 9; Length 153;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 33 LAGVIV 48
    11111
DB 120 LAGVIV 125

```

```

RESULT 43
US-09-825-301-27
: Sequence 27, Application US/09825301
: Patent No. US20020009738A1
: GENERAL INFORMATION:
: APPLICANT: Houghton, Raymond L.
: APPLICANT: Dillon, David C.
: APPLICANT: Molesh, David A.
: APPLICANT: Xu, Jiaoguo
: APPLICANT: Zehender, Barbara
: APPLICANT: Persing, David H.
: TITLE OF INVENTION: METHODS, COMPOSITIONS AND KITS FOR THE DETECTION
: FILE REFERENCE: 210121.513
: CURRENT APPLICATION NUMBER: US/09/825,301
: PRIOR FILING DATE: 2001-04-02
: NUMBER OF SEQ ID NOS: 77

```



```

: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 27
: LENGTH: 154
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (148)
: OTHER INFORMATION: Xaa = Any Amino Acid
US-09-825-301-27

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 154;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LTRKA 162
DB 33 LTRKA 38

RESULT 44
US-09-604-287A-471
: Sequence 471, Application US/09604287A
: Patent No. US20020064872A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Dillon, David C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepler, William T.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.470C7
: CURRENT APPLICATION NUMBER: US/09/604.287A
: CURRENT FILING DATE: 2000-06-22
: NUMBER OF SEQ ID NOS: 489
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 471
: LENGTH: 154
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (148)
: OTHER INFORMATION: Xaa = Any Amino Acid
US-09-604-287A-471

Query Match
Best Local Similarity 100.0%; Score 6; DB 10; Length 154;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LTRKA 162
DB 33 LTRKA 38

RESULT 45
US-10-007-805-471
: Sequence 471, Application US/10007805
: Patent No. US20020150581A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Dillon, David C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepler, William T.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedvick, Thomas S.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Durham, Margarita

```

```

: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.470C10
: CURRENT APPLICATION NUMBER: US/10/007.805
: CURRENT FILING DATE: 2001-12-07
: NUMBER OF SEQ ID NOS: 593
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 471
: LENGTH: 154
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: 148
: OTHER INFORMATION: Xaa = Any Amino Acid
US-10-007-805-471

Query Match
Best Local Similarity 100.0%; Score 6; DB 12; Length 154;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 157 LTRKA 162
DB 33 LTRKA 38

```

Search completed: December 7, 2002, 11:18:48
Job time : 18 secs

The first part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present. The second part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present.

TITLES

JOURNAL OF
MEDICAL NURSING

COMMENT

Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X., Hu, G., Lu, J., Chen, Y. and Han, Z. Insight into hepatocellular carcinogenesis at transcriptional level by comparing gene expression profiles of hepatocellular carcinoma with those of corresponding noncancerous liver Proc Natl Acad Sci U S A 98 (26), 15089-15094 (2001)

21625106

Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

FEATURES

BASE COUNT	168 a	198 c	176 g	162 t	2 others
ORIGIN					

Query Match	25.0%	Score 385;	DB 10;	Length 706;
Best Local Similarity	100.0%	Pred. No. 6.8e+183;		
Matches 385:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	830	GAAGCGCTGGCGCCGCACTGTGCTTCCCAAGATGAGCATGCTTCCCAAGSAAAGCTTCTCTG	889
Db	145	GAAGCCCTGTGGCCGCCACTGTCTCCCAAGSAGATGACATCTCTTCCCAAGSAAAGCTTCTACTCT	204
QY	890	ATGTCTACTGCCGAGCGSAACTGGACACGACTTGGTCAACGCGCTGGCCAGSAGTGAAGGCGCT	949
Db	205	ATGTCTACTGTCTGAATGAGAACTGGTCAAGATCTGGTCAAGCTGGTGAAGSAGTGAAGGCGCT	264
QY	950	CAGCTCTGTCTAATCAAAACTGCTGAGGAGTACAACTTCTTATAGCTTGCAGACTTCCAGG	1007
Db	265	CAGCTCTGTCTAATCAAAACTCTTGAAGATCTGAACACTTCTTATAGCTTGCAGACTTCCAGG	324
QY	1010	AGTAAAGCGCTTCTGCTGGATGGGAGCTTTCAGACTTAATATAGTAAGGCTAAGTGGCTAATGG	1066
Db	325	AGTAAAGCGCTTCTGCTGGATGGGAGCTTTCAGACTTAATATAGTAAGGCTAAGTGGCTAATGG	384
QY	1070	GTCGAGCGCTACCTCTGTGCACCGACGCTTCCAGCGTACTGGAGCAAGTGGAGAGATCCCAAC	1127
Db	385	GTCGAGCGCTACCTCTGTGCACCAAGCTTTCAGCGGTACTGGAGCAAGTGGAGAGATCCCAAC	444
QY	1130	AATAGCGGCAATGAAGATCTGTGCTGAATTTAGTGGCAAGTGGAGAGCAACAATCGATCT	1187
Db	445	AATAGCGGCAATGAAGACTGTGGCGAATTTAGTGGCAAGTGGAGAGCAACAATCGATCT	504
QY	1190	GAGCTTGACAATTACTGGATCTGCA	1214
Db	505	GAGCTTGACAATTACTGGATCTGCA	529

RESULT 4	BE562903
LOCUS	BE562903
DEFINITION	6013363329P1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3649319-5',
ACCESSION	BE562903
VERSION	BE562903.1 GI:9806623
KEYWORDS	EST
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1. (bases 1 to 828)	NIH-M&E	http://mre.cdc.nih.gov/	National Institutes of Health, Mammalian Gene Collection	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.			

FEATURES

Sources

```

/organism "Homo sapiens"
/db_xref "taxon:9606"
/cclone="IMAGE:3640319"
/cclone_1fb "NH_M9C_44"
/tissue_type="endometrium, adenocarcinoma"
/lab_host="JH108 (phage-resistant)"
/notes="Oradax intrus; Vector: pORF7; Site: EcoRI; cDNA made by oligo-dT priming; cloned into EcoRI/XhoI sites using the pAdaptor: ctk/ATCAG(S). Library constructed in the laboratory of Gerald M. Rubin (UC, California), Berkeley) using Zap-cDNA synthesis (Stratagene) and Superscript II RT (Life

```

Query Match	24.5%	Score 378;	EB 10;	Length:
Best Local Similarity	100.0%;	Prod. No. 2.3e-179;		
Matches 378;	Conservative	0;	Mismatches	0;
			Indel.	0;

QY	397	AGATTCAAGAGTGGGTAGCTGTTTAACTGATCAATCAATTCGAGCAATATG	100	414
Db	117	AGGTGCAAGGTGCTCCAGCTGCTTAAGTTCAGAGAAATATCGAGCAAAAGC	100	216
QY	457	GAACCTGACDCCAGCTTAAAGTGGCAATGGGTAAGTCTGACAGAAATCCAG	100	510
Db	177	GAACCTGACCCAGCTTTAAAGTGGCAATGGGTAAGCTCTGCAAGAAATCGAG	100	216
QY	517	GATCTACCAAGAGCTACCTAGTGAATGAAGCTGATGAGTTGGTATCT	100	574
Db	237	GATCTACCAAGAGCTAGCTAGTGAATGAAGCTGATGAGTTGGTATCT	100	216
QY	577	GTGTCAGAGATGATTAACAGAGATTAATGATGGGTGAAAGTGGCAATGGGTGA	100	634
Db	297	GTGTCAGAGATGATTAACAGAGATTAATGATGGGTGAAAGTGGCAATGGGTGA	100	216
QY	637	GAATTCGAAGCTGCAGAGATTTACATCAAGCTGACGTCGAGCTGAAGCTG	100	696
Db	357	GAATTCGAAGCTGCAGAGATTTATATGAGAGTGGAGAGAGCTGAAAGCTG	100	416
QY	697	GTTGCCAGAGAAATCGCAAGTGTGAGTAATCTACCGAGAGCTGATCTGGT	100	754
Db	417	GTTGCCAGAGAAATCCAAAGTGTGAGTAATCTACCGAGAGAGCTGATCTGGT	100	416
QY	757	AGTGGGTGAGTTGCCAGA	774	
Db	477	AGTGGGTGAGTTGCCAGA	494	

RESULT	LOCUS	DEFINITION	ACCESSION
5			
B1827055	B1827055	786 bp, mRNA	
	603075726F1 NIH_MGC_119	Homo sapiens cDNA clone	
		mRNA sequence.	
	B1827055		

[illegible]

	RESULT	6
R98113	LOCUS	503 bp mRNA linear EST J1-SEP-1995
	DEFINITION	Human spleen INFIS Homo sapiens cDNA clone IMAGE7445; similar to SH45274 AB0274 HIV G120-PRIME; NC-CATHE ELECTIN - 1, mRNA sequence.
	ACCESSION	R98113
	VERSION	R98113.1 GI:983773
	KEYWORDS	EST
	SOURCE	human:
	ORGANISM	Homo sapiens
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;
	REFERENCE	1 (bases 1 to 503)
	AUTHORS	Miller,L., Clark,N., Dubnue,T., Elliston,K., Hawkins,M., Holman,J.M., Holtzman,M., Kuchta,T., Le,M., Lennon,G., Marra,M., Parsons,J., Pfeiffer,E., Wolfberg,T., Soares,M., Tan,P., Trivaskis,F., Waterson,R., Williamson,A., Woldmann,P. and Wilson,R.
	TITLE	The WashU Merck EST Project
	JOURNAL	unpublished (1995)
	COMMENT	Contact: Wilson RK Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Insert Size: 1845 High quality sequence stops: 370 Source: IMAG Consortium, LINC. This clone is available royalty-free through LINC : contact the IMAG Consortium (info@image.linc.gov) for further information. Insert Length: 1845 Std Error: 0.00 Seq primer: M13Rp1 High quality sequence stop: 370.
FEATURES	Location/Qualifiers	
source	1..503	
	/organism="Homo sapiens"	
	/db_xref="GDB:3769764"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:200714"	
	/clone_lib="States fetal liver spleen INFIS"	
	/sex="male"	
	/day_stage="20 week post conception fetus"	
	/lab_host="DH10B (ampicillin resistant)"	
	/note="origin: Liver and Spleen, Vector: pTZ19 (Pharmacia), mate:"origin: Liver and Spleen, Vector: pTZ19 (Pharmacia), 1st strand cDNA was fitted with a Pac I site_2, Eco RI, 15' AACTCGAACAATTAAAGATCTTTTCTTTTTTTT 3'), double stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTZ19 vector. Library went through one round of normalization. Library constructed by Beate Soares and Matthew Bonaldo."	
BASE COUNT	126 a 137 c 128 g 105 t	7 others
ORIGIN		
Query Match	20.9%, Score 322, DB 14; Length 503;	
Best Local Similarity	100.0%; Pred No 4,3e-151;	
Matches	322, Conservative 0, Mismatches 0; Indels 0; Gaps 0;	
126	GAAATCTCTGGGTCAACTGCTGTAAGGCATCGAATTCTTGTAAGCAAACTGTAATGCT	889
127	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	111
128	CCAAAGCCTCTGGGTCAACTGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	125
129	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	158
130	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	194
131	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	227
132	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	264
133	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	297
134	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	334
135	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	367
136	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	404
137	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	437
138	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	474
139	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	507
140	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	544
141	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	577
142	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	614
143	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	647
144	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	684
145	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	717
146	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	754
147	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	787
148	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	824
149	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	857
150	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	894
151	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	927
152	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	964
153	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	997
154	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	1034
155	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1067
156	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	1104
157	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1137
158	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	1174
159	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1207
160	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	1244
161	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1277
162	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	1314
163	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1347
164	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	1384
165	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1417
166	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	1454
167	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1487
168	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	1524
169	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1557
170	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT	1594
171	TTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT	1627
172	ATCTCTGAACTGTAATGCTGTAAGGCATCGAATTCTTGTAAGCAAACTTACTGTT</	

QY	1010	AGTAAACGGCTTCCTGCGATGGATTTCAATCATAAATCATGAAGAGTAGGTGTAATG	1069
DB	246	AGTAAACGGCTTCCTGCGATGGATTTCAATCATAAATCATGAAGAGTAGGTGTAATG	305
QY	1070	GTCAGACGGCTACCTTGTCATCTTCACTTCCAGSDDGTAAGTGGAGAACGGCAAC	1129
DB	306	GTCAGACGGCTACCTTGTCATCTTCACTTCCAGSDDGTAAGTGGAGAACGGCAAC	365
QY	1130	AATACGGGCAATGACACTGTG	1151
DB	466	AATACGGGCAATGACACTGTG	487
RESULT 7			
AU140165			
LOCUS			
DEFINITION	AU140165 PLACE2 Homo sapiens cDNA clone PLACE23900662.5; mRNA		
ACCESSION	AU140165		
VERSION	AU140165.1 GI:11001686		
KEYWORDS	EST.		
SOURCE	HST.		
ORGANISM	human.		
	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo		
TITLE	1 (bases 1 to 792)		
	Ota,T., Suzuki,Y., Saito,K., Ishii,S., Yamamoto,I., Sugiyama,T.,		
	Nishikawa,T., Nakamura,Y., Sudano,S., Masuko,Y. and Isogai,T.		
	HRI human cDNA project (Ota,T., Suzuki,Y., Saito,K., Ishii,S.,		
	Yamamoto,J., Sugiyama,T., Nishikawa,T., Nakamura,Y., Sudano,S.,		
	Masuko,Y., Isogai,T.)		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Isogai		
	Genomics Laboratory		
	Helix Research Institute		
	1532-3 Yana, Kisarazu, Chiba 292-0412, Japan		
	Tel.: 81-438-52-3975		
	Fax: 81-438-52-3986		
	Email: genomics@hri.co.jp		
	HRI human cDNA project; 5'- & 3'-end one pass sequence ng; Helix		
	Research Institute; cDNA library construction; Department of		
	Virology, Institute of Medical Science, University of Tokyo, and		
	Helix Research Institute.		
FEATURES			
Source	Location/Qualifiers		
	1..792		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="PLACE2000062"		
	/clone_lib="PLACE2"		
	/tissue_type="Placenta"		
	/note="Vector: pMR1BSFL3"		
BASE COUNT	214 a 191 c 234 g 149 t	4 others	
ORIGIN			
Query Match	17.8%; Score 275; DB 9; Length 792;		
Best Local Similarity	99.5%; Pred. No. 2; e-127;		
Matches	375; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY	398	GTCGCCAGGTCGCCAGCTCCTTAAGTCAGACAACATCCAGCACAGACGCATCATACACAG	457
DB	53	GTCGCCAGGTCGCCAGCTCCTTAAGTCAGACAACATCCAGCACAGACGCATCATACACAG	112
QY	458	AACGTGACCAGCTTAACTGCTACCTGGGTGAGCTGTACAGAGAATATCCAGCTGCCAGGAG	517
DB	113	AACGTGACCAGCTTAACTGCTACCTGGGTGAGCTGTACAGAGAATATCCAGCTGCCAGGAG	172
QY	518	ATTACCAAGAGCTGACTAGCTAGCTGAAAGGCTGAGTGGGTGAGTGGCCAGAGAATTCAG	577
DB	173	ATTACCAAGAGCTGACTAGCTGAGGCTGAAAGGCTGAGTGGGTGAGTGGCCAGAGAATTCAG	232
QY	578	CTGACAGAGATCTACCAAGAGCTGACCGGCTTGAAAGGCTGTGAGTGGGTGAGTGGCCAGAG	637
DB	233	CTGACAGAGATCTACCAAGAGCTGACCGGCTTGAAAGGCTGTGAGTGGGTGAGTGGCCAGAG	292

QY	1302	CTACACCGGTGACCCCTTTTGGACTGGGAGATTTGGTGGTTGACAGACCTT
QY	638	AAATCCAAAGCTGTAGAGGAGATCTATCTAGAGAGAGTGTGACCTGGAGGCTGTGG
DB	293	AAATCCAAAGCTACAGAGGAGATCTATCTAGAGAGTGTGACCTGGAGGCTGTGG
QY	698	TTGCTCAGAGAAATCCAAAGTGTGGAGATCTCTAGAGAGTGTGAGATCTGAG
DB	353	TTTCTCAGAGAAATCCAAAGTGTGGAGATCTCTAGAGAGTGTGAGATCTGAG
QY	758	GTGGATGAGTTGGCAGG 774
DB	413	GTGGATGAGTTGGCAGG 429

RESULT 8
LOCUS AA779963 429 bp mRNA linear
DEFINITION 2188c08.s1 Soares fetal liver spleen INFILS.S1 H
ACCESSION AA779963
VERSION AA779963
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Gracilata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoi
AUTHORS 1 (bases 1 to 326)
Hillier, L., Allen, M., Bowles, L., Dubugue, T., Golt
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G.
J., Moore, B., Scheinberg, K., Steptoe, M., Tan
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST project
Unpublished (1997)
TITLE Contact: Wilson R.
JOURNAL Washington University School of Medicine
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, M
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LIA
IMAGE Consortium (info:lmage.lim.gov) for LITR.
Seq primer: -40m3 fwd. Et from Amersham
High quality sequence stop: 246.
FEATURES
source Location/Qualifiers
1..326
/organism="Homo sapiens"
/db_xref="Gene:3752980"
/db_xref="taxon:9606"
/clone="IMAGE:461966"
/clone_lib="Soares fetal_liver_spleen_15"
/sex="male"
/seq_start="20 week post conception for
/label_host="Human (amplification resistant)
/note="Organ: Liver and Spleen, Vector
with a modified polylinker Site_1; Pac
This is a subtracted version of the o
liver spleen INFILS library. 1st stran
with a Pac 1 - oligo(dT) primer (5'
AACTCGAAGAAATTAATTAAGATCTTTTTTTTTTT
double-stranded cDNA was ligated to Eco
(Pharmacia), digested with Pac 1 and c
and Eco RI sites of the modified pT777
went through one round of normalization
constructed by Bento Soares and M.Pati
BASE COUNT 93 a 64 c 87 g 82 t
ORIGIN
Query Match 15.6%; Score 241; DB 9; Length
Best Local Similarity 100.0%; Fred. No. 3 Re-110.
Matches 241; Conservative 0; Mismatches 0; Indel

Db 245 CTACACCCCGGTGACCTTTTGAAGTGGTCTTGGTTGAGGAGGCTGATCTTTCAGG 186
 QY 1362 CTGCAAGTAAATAGGAAATTAATCTTCCAGTCAACCAATGTAATGACATAGGCTATGGA 1421
 Db 185 CTGCAAGTAAATAGGAAATTAATCTTCCAGTCAACCAATGTAATGACATAGGCTATGGA 125
 QY 1422 GGGCTGCAATACAACTCTTCCAGTCAACCAATGTAATGACATAGGCTATGGA 1481
 Db 125 GGGCTGCAATACAACTCTTCCAGTCAACCAATGTAATGACATAGGCTATGGA 66
 QY 1482 GGGCTGCTGAGTTGAGCTTTATATTAATGTAATGTAATGTTTACTGCAAAAAA 1541
 Db 65 GGGCTGCTGAGTTGAGCTTTATATTAATGTAATGTAATGTTTACTGCAAAAAA 6
 QY 1542 A 1542
 Db 5 A 5

RESULT 9

B0008265/c

LOCUS B0008265 438 bp. mRNA linear. EST 26-MAR-2002
 DEFINITION UT-H-E10-ayg-e-13-0-01-s1 NCL CGAP_E10 Homo sapiens cDNA clone

IMAGE: 5848424 3', mRNA sequence.

ACCESSION B0008265 GI:19733166

VERSION B0008265.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 438)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo
 TITLE NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-f@mail.nih.gov

FEATURES

location/Qualifiers
 1..438
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5838324"
 /clone_lib="NCI CGAP_E10"
 /issue_type="Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia)
 with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
 NCI CGAP_E10 is a cDNA library containing the following
 tissue(s): Chondrosarcoma. The library was constructed
 according to Bonaldi, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested
 with Not I, and cloned directionally into pT73-Pac
 vector. The oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dT)18 tail. The
 sequence tag for this library is AACTTCACAC.
 TAG_L1B-01-H-E10
 TAG_TISSUE=chondrosarcoma
 TAG_SEQ=AACTTCACAC"

BASE COUNT 130 a 96 c 102 g 105 t
 ORIGIN

Query Match 15.6%; Score 241; DB 14; Length 438;
 Best Local Similarity 100.0%; Pred. No. 3,20-110;

Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 CTACATGCTGAGTCTGAGCTTTATATTAATGTAATGTAATGTTTACTGCAAAAAA 1541
 Db 250 CTACATGCTGAGTCTGAGCTTTATATTAATGTAATGTAATGTTTACTGCAAAAAA 191

QY 1362 CTGCAAGTAAATAGGAAATTAATCTTCCAGTCAACCAATGTAATGACATAGGCTATGGA 1421
 Db 190 CTGCAAGTAAATAGGAAATTAATCTTCCAGTCAACCAATGTAATGACATAGGCTATGGA 131

QY 1422 GGGCTGCAATACAACTCTTCCAGTCAACCAATGTAATGACATAGGCTATGGA 1481
 Db 130 GGGCTGCAATACAACTCTTCCAGTCAACCAATGTAATGACATAGGCTATGGA 71

QY 1482 GGGCTGCTGAGTTGAGCTTTATATTAATGTAATGTAATGTTTACTGCAAAAAA 1541
 Db 70 GGGCTGCTGAGTTGAGCTTTATATTAATGTAATGTAATGTTTACTGCAAAAAA 11

QY 1542 A 1542
 Db 10 A 10

RESULT 10

B0006101/c

LOCUS B0006101 656 bp. mRNA linear. EST 26-MAR-2002
 DEFINITION UT-H-E11-ayg-n-09-0-01-s1 NCL CGAP_E11 Homo sapiens cDNA clone

IMAGE: 5845832 3', mRNA sequence.

ACCESSION B0006101 GI:19731001

VERSION B0006101.1

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 656)
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
 JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

COMMENT Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-f@mail.nih.gov

FEATURES

location/Qualifiers
 1..656
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5845832"
 /clone_lib="NCI CGAP_E11"
 /issue_type="Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia)
 with a modified polylinker; Site_1: EcoR I; Site_2: Not I;
 NCI CGAP_E11 is a normalized cDNA library containing the
 following tissue(s): Chondrosarcoma. The library was
 constructed according to Bonaldi, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is

Query Match	9.38	Score 144	nr	Length 455
Best Local Similarity	100.0%	Prod No. 2.7e e1		
Matches 144:	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0
OY	1484	TCGCCGAGTACACCATGGCATCCAGACAGGCGCATGAGAGCTGCATATTAACCTCCTCCAC	1443	
Db	150	TCGCCGAGTACACCATGGCATCCAGACAGGCGCATGAGAGCTGCATATTAACCTCCTCCAC	91	
OY	1444	CCCCACATCTTTCTTTGTCTCTATACATGCTTCACATTGGCGTGTTCGAGTTGAGCT	1503	
Db	90	CCCCACATCTTTCTTTGTCTCTATACATGCTTCACATTGGCGTGTTCGAGTTGAGCT	31	
OY	1504	TTATAATAAGTGTAAATCTGT	1527	
Db	30	TTATAATAAGTGTAAATCTGT	7	

RESULT 19	LOCUS	DEFINITION	ACCSSION
H00185/..	H00185	302 bp mRNA linear	F0119-10q
		Y122H03.s1 Soares placentu N02HF Homo sapiens cDNA clone IMAGE:149525 3', mRNA sequence.	

ACCELERATION	000105	
VERSION	H00185.1	G1:863118
KEYWORDS	FST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo
 1 (passes 1 to 302)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman,
 M., Hultman, M., Kuchta, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
 Ritzkin, L., Rohlfing, T., Soares, M., Tan, F., Tevaskis, F., Waterston
 R., Williamson, A., Wohldmann, P. and Wilson, R.
 The WashU-Merck EST project
 Unpublished (1995)
 Contact: Willson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63103
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewartson.wustl.edu
Insert Size: 689
High quality sequence stops: 217
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL. Contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 689 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 217.

FEATURES	Location/Qualifiers
source	1. .302

```

/organism="Homo sapiens"
/db_xref="GPR:561404"
/db_xref="Taxon:9606"
/clone="IMAGE:149525"
/clone.lib="Soares placenta Nb2Hp"
/sex="Female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ. placenta: Vector: pT73b (Pharmacia) with a
modified polylinker. Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
AATCGAATATTTGCTGCTGACGATTTTTTTTTTTTTTTT 3']
double-stranded cDNA was ligated to Eco RI receptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bonaldo. "

```

ORIGIN	9, 38;	Score 143;	DB 14;	Length 302;
Query Match				

	Best Local Similarity	100.0%	Fract. No.	9.26-61				
	Matches	143	Conservative	0	Mismatch	0	Indels	0
QY	1384	TCGCGGACGTCAGGCAAGATGCGATGTCATATGATAGGGGATGGAGGATCCGATAAAGC						
Db	143							
		TCGCGGACGTCAGGCAAGATGCGATGTCATATGATAGGGGATGGAGGATCCGATAAAGC						
QY	1444	CCGCGACATCTTCTCTTGTGCTGATATATATGTCCTTCGACATTTGGCTGTTTCTGAG						
Db	83	CCGCGACATCTTCTCTTGTGCTGATATATATGTCCTTCGACATTTGGCTGTTTCTGAG						
QY	1504	TTATATTAAGCGTAAATGTTG	1526					
Db	23	TTATATTAAGCGTAAATGTTG	1					

RESULT 20
AV645629

LOCUS	AV645629	459 bp	mRNA	100%
DEFINITION	AV645629	GLC	HOMO SAPIENS	GLNA
ACCESSION	AV645629			

VERSION	AV645629.1	GI:98676446
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata
 Mammalia; Eutheria; Primates; Catarrhini; Homiini
 1 (bases 1 to 359)
 REFERENCE
 Xu, X., Huang, T., Yu, Z., Li, P., Wu, Z., Yao, C.,

TITLE Xiao, H., Wu, J., Liu, F., Huang, Y., Cheng, Z., Li, N., Shen, K., Lu, G., Fu, G., Zhou, M., Xu, S., Gu, W., Fu, H., G., Cai, J., Chen, Y. and Han, Z.
Insight into hepatocellular carcinogenesis at the

by comparing gene expression profiles of hepatocarcinoma with those of corresponding noncancerous liver

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094
2001

RECEIVED	DATE	COMMENT
2102/2100		Contact: Zequang Han Chinese National Genome Center at Shanghai 351 Guo Shoujing Road, Zhongjiang Hi-Tech Park, i 201202 P. China

201203, P. R. CHINA
Tel: 86-21-50801919 (ext. 45)
Fax: 86-21-50801922
Email: hanzy@hgc.sh.cn

FEATURES	source
This clone is available at CHMO in Shanghai.	Location/individuals
1. .359	/organism "Homo sapiens"

```

/db_xref "taxon: 3606"
/clone="G1C7C012"
/clone_lib="G1C"
/tissue_type "Corresponding non cancerous"

```

```

/dev_stage="Adult"
/Lab_host="SoiR"
/node="Vector: phagescript sk(-); Site:
  pho1"

```

BASE COUNT	85 a	111 c	91 d	72 t
ORIGIN				

Query match	6.8%	Score 152	DB 102	Length
Best Similarity	103.03	Prod. No.	9.7e-57	
Best Local	135	Conservative	0	Mismatches
			0	Indels

830 GAAAGCAGTGGGAGCAGTGGTAAAGACTGGATCTTCCAAAGGAA
 |||||
 145 GAAGCCTGGAGCCACATGCTGTAAGACTGGATCTCTTCCAAAGGAA *

QY	890 ATGTTACTCTCCACGGCAATTCGTAACTAATCGTGTAACGCTGCACAATA
Ddb	205 ATGTCTAAATCCCAGCAGAAATCGTAACTAACATCGTGAACGCTGTATTAAA

QY 950 CAGCTCGTCTATTC 964
|||||

```

Db      265  CACCTGCTGCTGCTATG 279

RESULT 21
LOCUS   AA448145
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE
1. 411
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:782758"
/clone.lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/notes="Vector: pTR30-Pac (Pharmacia) with a modified EcoRI linker, Site_1 Not I; Site_2, Eco RI, 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer (5'-TCCTCCCAATCTCAAGATGCGGCGGCGGCGGCGCAATTTTTTTTTTTT-3'). Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTR3 vector. Library went through one round of normalization to Col5, and was constructed by Puro Soares and M Fatima Formado."
BASE COUNT      86 a      103 c      100 g      122 t
ORIGIN
      1
      2
      3
      4
      5
      6
      7
      8
      9
     10
     11
     12
     13
     14
     15
     16
     17
     18
     19
     20
     21
     22
     23
     24
     25
     26
     27
     28
     29
     30
     31
     32
     33
     34
     35
     36
     37
     38
     39
     40
     41
     42
     43
     44
     45
     46
     47
     48
     49
     50
     51
     52
     53
     54
     55
     56
     57
     58
     59
     60
     61
     62
     63
     64
     65
     66
     67
     68
     69
     70
     71
     72
     73
     74
     75
     76
     77
     78
     79
     80
     81
     82
     83
     84
     85
     86
     87
     88
     89
     90
     91
     92
     93
     94
     95
     96
     97
     98
     99
    100
    101
    102
    103
    104
    105
    106
    107
    108
    109
    110
    111
    112
    113
    114
    115
    116
    117
    118
    119
    120
    121
    122
    123
    124
    125
    126
    127
    128
    129
    130
    131
    132
    133
    134
    135
    136
    137
    138
    139
    140
    141
    142
    143
    144
    145
    146
    147
    148
    149
    150
    151
    152
    153
    154
    155
    156
    157
    158
    159
    160
    161
    162
    163
    164
    165
    166
    167
    168
    169
    170
    171
    172
    173
    174
    175
    176
    177
    178
    179
    180
    181
    182
    183
    184
    185
    186
    187
    188
    189
    190
    191
    192
    193
    194
    195
    196
    197
    198
    199
    200
    201
    202
    203
    204
    205
    206
    207
    208
    209
    210
    211
    212
    213
    214
    215
    216
    217
    218
    219
    220
    221
    222
    223
    224
    225
    226
    227
    228
    229
    230
    231
    232
    233
    234
    235
    236
    237
    238
    239
    240
    241
    242
    243
    244
    245
    246
    247
    248
    249
    250
    251
    252
    253
    254
    255
    256
    257
    258
    259
    260
    261
    262
    263
    264
    265
    266
    267
    268
    269
    270
    271
    272
    273
    274
    275
    276
    277
    278
    279
    280
    281
    282
    283
    284
    285
    286
    287
    288
    289
    290
    291
    292
    293
    294
    295
    296
    297
    298
    299
    300
    301
    302
    303
    304
    305
    306
    307
    308
    309
    310
    311
    312
    313
    314
    315
    316
    317
    318
    319
    320
    321
    322
    323
    324
    325
    326
    327
    328
    329
    330
    331
    332
    333
    334
    335
    336
    337
    338
    339
    340
    341
    342
    343
    344
    345
    346
    347
    348
    349
    350
    351
    352
    353
    354
    355
    356
    357
    358
    359
    360
    361
    362
    363
    364
    365
    366
    367
    368
    369
    370
    371
    372
    373
    374
    375
    376
    377
    378
    379
    380
    381
    382
    383
    384
    385
    386
    387
    388
    389
    390
    391
    392
    393
    394
    395
    396
    397
    398
    399
    400
    401
    402
    403
    404
    405
    406
    407
    408
    409
    410
    411
    412
    413
    414
    415
    416
    417
    418
    419
    420
    421
    422
    423
    424
    425
    426
    427
    428
    429
    430
    431
    432
    433
    434
    435
    436
    437
    438
    439
    440
    441
    442
    443
    444
    445
    446
    447
    448
    449
    450
    451
    452
    453
    454
    455
    456
    457
    458
    459
    460
    461
    462
    463
    464
    465
    466
    467
    468
    469
    470
    471
    472
    473
    474
    475
    476
    477
    478
    479
    480
    481
    482
    483
    484
    485
    486
    487
    488
    489
    490
    491
    492
    493
    494
    495
    496
    497
    498
    499
    500
    501
    502
    503
    504
    505
    506
    507
    508
    509
    510
    511
    512
    513
    514
    515
    516
    517
    518
    519
    520
    521
    522
    523
    524
    525
    526
    527
    528
    529
    530
    531
    532
    533
    534
    535
    536
    537
    538
    539
    540
    541
    542
    543
    544
    545
    546
    547
    548
    549
    550
    551
    552
    553
    554
    555
    556
    557
    558
    559
    560
    561
    562
    563
    564
    565
    566
    567
    568
    569
    570
    571
    572
    573
    574
    575
    576
    577
    578
    579
    580
    581
    582
    583
    584
    585
    586
    587
    588
    589
    590
    591
    592
    593
    594
    595
    596
    597
    598
    599
    600
    601
    602
    603
    604
    605
    606
    607
    608
    609
    610
    611
    612
    613
    614
    615
    616
    617
    618
    619
    620
    621
    622
    623
    624
    625
    626
    627
    628
    629
    630
    631
    632
    633
    634
    635
    636
    637
    638
    639
    640
    641
    642
    643
    
```

[illegible]

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota, Metazoa, Chordata, Claviata, Vertebrata: Eucleustomi; Mammalia, Eutheria: Primates; Catarrhini: Hominoidea: Homo. 1 (bases 1 to 499)
AUTHORS	Hillier, L., Allen, M., Bowles, L., Dubouque, T., Geisler, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenger, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
TITLE	WashU-Merck EST Project 1997
JOURNAL	Unpublished (1997)
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -28m3 rev2 ET from Amersham High quality sequence stop: 495. Location/Qualifiers
FEATURES	1..499 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:782332" /clone_lib="Soares_testis_NHT" /sex="male" /lab_host="DH10H" /note="Vector: pT730-Pac (Pharmacia) with a modified polylinker. Site_1: Not 1; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories , Inc., and primed with a Not I - oligo(dT) primer [5', TCTTACCACTCGAAGTGGAGCGCGCCGCAATTTTCTTTTCTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cct5, and was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT	94 a 135 c 112 g 158 t
ORIGIN	
Query Match	6.7%; Score 104; DB 9; Length 499;
Match Local Similarity	100.0%; Pred. NO. 4; 1e-41;
Matches 104; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1198	CAATTACTGCATGTGCAGAAACCCGCGACACCTGCTTAGAGACGAAATAGTGTTCCTG 1257
DB 1	CAATTACTGCATGTGCAGAAACCCGCGACACCTGCTTAGAGACGAAATAGTGTTCCTG 60
OY 1258	CTAGCCCTACGCTTCATCTGTGCTATAGACGAGCACTTACCCACTT 1301
DB 61	CTAGCCCTACGCTTCATCTGTGCTATAGACGAGCACTTACCCACTT 104
RESULT 24	
LOCUS	H009487 393 bp mRNA linear EST 19-JUN-1995
DEFINITION	yj28b04.s1 Soares placenta Ni2HP Homo sapiens cDNA clone
ACCESSION	IMAGE:150055.3, mRNA sequence.
VERSION	H009487
KEYWORDS	H009487.1 GI:863920
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota, Metazoa, Chordata, Claviata, Vertebrata: Euteleostomi; Mammalia: Eutheria: Primates; Catarrhini: Hominoidea: Homo. 1 (bases 1 to 393) Hillier, L., Clark, N., Dubouque, T., Ellison, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Roeding, T., Soares, M., Tan, F., Treaskis, S., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
TITLE	The WashU-Merck EST Project

JOURNAL COMMENT		Unpublished (1995)	
Contact: Wilson RK		Washington University School of Medicine	
4444 Forest Park Parkway, Box 8501, St. Louis, MO			
Tel: 314 286 1800			
Fax: 314 286 1810			
Email: est.watson.wustl@edn			
Insert Size: 662			
High quality sequence steps: 297			
Source: IMAGE Consortium, LBNL			
This clone is available royalty-free through LBNL			
IMAGE Consortium (info: image.lbnl.gov) for further			
Insert Length: 662 Std Error: 0.00			
Seq primer: Promega -21mt13			
High quality sequence step: 297.			
Location/Qualifiers			
1..493			
/organism="Homo sapiens"			
/db_xref="db:561999"			
/db_xref="taxon:9606"			
/clone="IMAGE:150055"			
/clone_lib="Soares placenta ND2HP"			
/sex="female"			
/dev_stage="placenta obtained at birth"			
/lab_host="pBlhis (ampicillin resistant)"			
/note="Organ: placenta; Vector: pTZ19"			
modified polylinker; Site: 1; Not 1; Site:			
strand cDNA was primed with a Not 1 site.			
AACGCAACAAATTCGCGCGGACAGAAATTTTCTT			
double-stranded cDNA was ligated to Eco			
(thamacia), digested with Not 1 and Eco			
and Eco RI sites of the modified pTZ19			
went through one round of normalization			
constructed by Bento Soares and M.Felipe			
BASE COUNT 111 a 84 c 103 g 84 t 11 other			
ORIGIN			
Query Match 6.08; Score 94; DB 14; Length			
Best Local Similarity 99.38; Pred. No. 156-35;			
Matches 143; Conservative 0; Mismatches 1; Indel:			
QY 1384	TCGCCGAGCAGCAACATAGGATGCAACAGAG333GAG5AGGCTGCATACAA		
DB 150	TCCTTAAATCAAAAGAGATGCAACAGAG333GAG5AGGCTGCATACAA		
QY 1444	CCCCACATCTTCTTGCTCTATACATCTCTGCATTTGGCTGTTTCGA		
DB 90	CCCCACATCTTCTTGCTCTATACATCTCTGCATTTGGCTGTTTCGA		
QY 1504	TATATAAAGTGTGTAATCTTCT 1527		
DB 30	TATATAAAGTGTGTAATCTTCT 7		
RESULT 25			
AA002266/c 428 bp mRNA 1100g MAY 1997			
LOCUS			
DEFINITION ZH82003.s1 Soares fetal liver spleen INFUS_S1 Ec			
cGene IMAGE:42781.c, mRNA source:ec			
ACCESSION AA002266			
VERSION AA002266.1 GI:1445181			
KEYWORDS EST.			
SOURCE human.			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Granulata; Vertebrata			
Mammalia; Eutheria; Primates; Catarrhini; Homin			
1 (bases 1 to 428)			
Biller, L., Lennon, G., Becker, M., Bonaldo, M.F.,			
Chissoe, S., Dietrich, N., Dubugue, J., Favello, A.,			
M., Hultman, M., Kuraba, I., Lacy, M., Le, M., Le, N.			
B., Morris, M., Parsons, J., Prade, C., Riklin, I.,			
Schellenberg, K., Soares, M.B., Tan, F., Thierry-M			
Underwood, K., Wohlmann, P., Waterston, R., Willet			

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer genome project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/getitem2.pl?i=552-PM2-TN0025-050
 900-007-003&i=2000-09-05&c4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 78
 High quality sequence stop: 254
 Location/Qualifiers
 1..254
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="TN0025"
 /dev_stage="Adult"
 /note="Organ: testis-normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from OKESHS PCR (U S letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
 59 a 67 c 56 g 72 t

Query Match 2.28; Score 34; EB 12; Length 254;
 Best Local Similarity 100.0%; Pred. No. 9; E=06;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1228 CTGCTTCAGACGACATAGTGTTCCTGCTAG 1261
 DB 94 CTGCTTCAGACGACATAGTGTTCCTGCTAG 127

RESULT 28
 AA918594 405 bp mRNA linear EST 10-JUN-1998
 LOCUS 0153c12.s1 Soares_NFL_T-GRC_S1 Homo sapiens cDNA clone
 DEFINITION IMAGE:1527190.3', mRNA sequence.
 AA918594
 VERSION AA918594.1 GI:3058484
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 405)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT unpublished (1997)
 Email: cgaps@mail.nih.gov
 This clone is available royalty-free through NCI; contact the
 IMAGE Consortium (info@image.jiml.gov) for further information.
 Insert Length: 513 Std Error: 0.00
 Seq primer: 40ml3 fwd. ET from Amersham
 High quality sequence stop: 338.
 Location/Qualifiers
 1..405
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1527190"
 /clone_lib="Soares_NFL_T-GRC_S1"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: p7770-Pac (Pharmacia) with
 a modified polylinker; Site 1: Not I; Site 2: Eco RI;
 Equal amounts of plasmid DNA from three normal
 libraries (testis lung NBH19W, testis NHT, and B-cell
 NCI-CGAP-GCH) were mixed, and ss circles were made in
 vitro. Following HAP purification, this DNA was used as
 tracer in a subtractive hybridization reaction. The direct

was PCR-amplified cDNAs from pools of
 from the same 3 libraries. The pools con-
 tain M.A.G.E. clones 297480, 620887, 682632,
 726408-728711, and 729096-731399. Submitted
 Soares and M. Fatima Honaido.
 BASE COUNT 100 a 88 c 89 g 128 t

Query Match 2.28; Score 34; EB 3; Length 4
 Best Local Similarity 100.0%; Pred. No. 8; E=06;
 Matches 34; Conservative 0; Mismatches 0; Indels 0;
 OY 1479 TTGCGCTTTCTCAGTGTGACCTTATATATA 1512
 DB 289 TTGCGCTTTCTCAGTGTGACCTTATATATA 22

RESULT 29
 A1217955 453 bp mRNA linear EST
 LOCUS q152c6.x1 Soares_testis NHT Homo sapiens cDNA
 DEFINITION IMAGE:1753642.3', mRNA sequence.
 A1217955
 VERSION A1217955.1 GI:3797770
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
 REFERENCE 1 (bases 1 to 453)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy
 Project (CGAP), Tumor Gene Index
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT unpublished (1997)
 Email: cgaps@mail.nih.gov
 cDNA library preparation: M. Bento Soares, Ph.D.
 , Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome
 Clone distribution: NCI-CGAP clone distribution
 found through the M.A.G.E. Consortium (http://
 www.bio.jiml.gov/ftp/image/image.html)
 Insert Length: 548 Std Error: 0.00
 Seq primer: 40UP from Gibco.
 Location/Qualifiers
 1..453
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1753642"
 /clone_lib="Soares_testis_NHT"
 /seq="male"
 /lab_host="DH10B"
 /note="Vector: p7770-Pac (Pharmacia) v
 polylinker; Site 1: Not I; Site 2: Eco RI;
 Double-stranded cDNA was ligated to Eco
 (Pharmacia), digested with Not I and
 and Eco RI sites of the modified p773
 went through one round of normalization
 constructed by Bento Soares and M. Fat

BASE COUNT 112 a 97 c 104 g 140 t

Query Match 2.28; Score 34; EB 9; Length 4
 Best Local Similarity 100.0%; Pred. No. 8; E=06;
 Matches 34; Conservative 0; Mismatches 0; Indels 0;
 OY 1479 TTGCGCTTTCTCAGTGTGACCTTATATATA 1512
 DB 289 TTGCGCTTTCTCAGTGTGACCTTATATATA 22

RESULT 29
 A1217955 453 bp mRNA linear EST
 LOCUS q152c6.x1 Soares_testis NHT Homo sapiens cDNA
 DEFINITION IMAGE:1753642.3', mRNA sequence.
 A1217955
 VERSION A1217955.1 GI:3797770
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
 REFERENCE 1 (bases 1 to 453)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy
 Project (CGAP), Tumor Gene Index
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT unpublished (1997)
 Email: cgaps@mail.nih.gov
 cDNA library preparation: M. Bento Soares, Ph.D.
 , Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome
 Clone distribution: NCI-CGAP clone distribution
 found through the M.A.G.E. Consortium (http://
 www.bio.jiml.gov/ftp/image/image.html)
 Insert Length: 548 Std Error: 0.00
 Seq primer: 40UP from Gibco.
 Location/Qualifiers
 1..453
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1753642"
 /clone_lib="Soares_testis_NHT"
 /seq="male"
 /lab_host="DH10B"
 /note="Vector: p7770-Pac (Pharmacia) v
 polylinker; Site 1: Not I; Site 2: Eco RI;
 Double-stranded cDNA was ligated to Eco
 (Pharmacia), digested with Not I and
 and Eco RI sites of the modified p773
 went through one round of normalization
 constructed by Bento Soares and M. Fat

RESULT 30
AA669250 472 bp mRNA linear EST 20 NOV 1997
LOCUS acc17c10.s1 Stralagene ovary (#937217) Homo sapiens cDNA clone
DEFINITION IMAGE:856746.3 similar to catallus Alu repetitive element?, mRNA
sequence.

ACCESSION AA669250
VERSION 1 GI:2630749
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 472)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,
Kilman, D., Kudaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellenger, K., Shapiro, M., Tan, F., Theisling, R.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST project
Unpublished (1997)
JOURNAL Contact: Wilson, R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Possible reverts: 4.4e-6; 1.7e-7 not found
Seq primer: -40mt fwd, ET from Amersham.
Location/Qualifiers
1..472
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:856746"
/sex="female"
/dev_stage="49 year old"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ ovary, Vector: Plasmid SK, Site: 1. EcoRI
Site: 2. XbaI, cloned unidirectionally. Primer: oligo di-
total ovary tissue, normal, Caucasian. Accurate insert
size: 0.8 kb; Uni ZAP XR Vector; 5' adaptor sequence: 5'
CAATCGGCGACGAG 3' -37 adaptor sequence: 5'
CGGAGTTTCTTTTCTTTT 37"

BASE COUNT 124 a 110 c 93 g 145 t
ORIGIN

Query Match 1.78; Score 26; DB 9; Length 472;
Best local similarity 100.0%; Pred No 0.092;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0

CY 1474 TTCATTGGCTTTCTGCTGTA 1499
|||||
DB 219 TTCATTGGCTTTCTGCTGTA 194

RESULT 31
H00237 478 bp mRNA linear EST 19 JUN 1995
LOCUS yj22h03.f1 Soares placenta Nb2HP Homo sapiens cDNA clone
DEFINITION IMAGE:149525.5, mRNA sequence.

ACCESSION H00237
VERSION 1 GI:863170
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 478)
Hillier, L., Clark, N., Dubuque, T., Ellison, K., Haskins, M., Hohman

TITLE JOSEPHAL
COMMENT: Unpublished (1995)
Contact: Wilson, R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert size: 689
High quality sequence stops: 322
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 689 Std Error: 0.00
Seq primer: M13p1
High quality sequence stop: 322.
Location/Qualifiers
1..478
/organism="Homo sapiens"
/db_xref="taxon:9606"
/db_xref="IMAGE:149525"
/clone_lib="Soares placenta Nb2HP"
/sex="female"
/dev_stage="placenta obtained at birth (full term)"
/lab_host="H10b (ampicillin resistant)"
/note="Organ: placenta; Vector: pTZ19 (Pharmacia) with a
modified polylinker. Site: 1. Not I, Site: 2. Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer (5'
AAGCGAAGAAATTCGCGCCGACGAGATTTTCTTTTCTTTT 37),
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTZ19 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 90 a 126 c 116 g 145 t
ORIGIN

Query Match 1.63; Score 24; DB 14; Length 478;
Best local similarity 100.0%; Pred No. 0.94;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0

CY 1337 GCAAGCTGCAAGCAAGCAAGT 1380
|||||
DB 413 GCAAGCTGCAAGCAAGCAAGT 436

RESULT 32
A1614060 281 bp mRNA linear EST 21 APR 1999
LOCUS yj34j01.31 Soares mammary gland NRIMG Mus musculus cDNA clone
DEFINITION IMAGE:185333.5 similar to SW-K19c_PAT p10716 KUPFERER CELL
RECEPTOR, mRNA sequence.

ACCESSION A1614060
VERSION 1 GI:4623227
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 281)
NCI-CCAP http://www.ncbi.nlm.nih.gov/nciccap.
National Cancer Institute, Cancer Genome Anatomy Project (CCAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: ega@sl-shall.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
M1:507321

This read is a RESEQUENCE of a previously sequenced mouse clone
 This read has been vetted (found to hit its original self in the
 correct orientation)
 Seq primer: -40RP from Gibco
 High quality sequence stop: 281.

FEATURES

SOURCE

Location/Qualifiers
 1. 244

/organism="Mus musculus"
 /strain="C57Bl/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:863233"
 /clone_lib="Soares_mammary_gland_NBMG"
 /sex="male"
 /tissue_type="mammary gland"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Organ: mammary gland; Vector: pT710-P3c (Pharmacia)
) with a modified polylinker site; Not 1; Site-2; Eco
 RI; 1st strand cDNA was primed with a Not I - oligo(dT)
 primer (5'
 TGTACCAATCTGAGTGGAGCGCGCGCAATGTTTTTTTTTTTTTTTTTTT
 T 3'); double-stranded cDNA was ligated to Eco RI
 adaptors (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT713 vector.
 RNA provided by Dr. Minoru Ko, Wayne State Univ library
 constructed and normalized by Bento Soares and M. Fatima
 Bonaldo."

BASE COUNT

70 a 64 c 81 g 69 t

ORIGIN

Query Match 1.54; Score 23; DB 9; Length 284;
 Best Local Similarity 100.0%; Pred No. 3.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1076 GGCTACCTCTGTGACCGAGCTT 1098

Db 147 GGCTACCTCTGTGACCGAGCTT 169

RESULT 33

LOCUS

AA510960 398 bp mRNA linear EST 08-JUL-1997
 v34401.1 Soares_mammary_gland_NBMG Mus musculus cDNA clone

IMAGE:863233 5' similar to SW:KICP_PAT P10716 RUPFERF CELL
 RECEPTOR mRNA sequence.

AA510960
 AA510960.1 GI:2248814

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

The Washu-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Maria M/Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.wustl.edu
 This clone is available royalty-free through LNC; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI-507421
 Seq primer: -28ml3 rev2 ET from Amerham
 High quality sequence stop: 318.
 Location/Qualifiers

source

1. 398
 /organism="Mus musculus"
 /strain="C57Bl/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:863233"
 /clone_lib="Soares_mammary_gland_NBMG"
 /sex="male"
 /tissue_type="mammary gland"
 /dev_stage="4 weeks"
 /lab_host="DH10B"
 /note="Organ: mammary gland; Vector: pT
) with a modified polylinker site; Not 1; Site-2; Eco
 RI; 1st strand cDNA was primed with a Not
 primer (5'
 TGTACCAATCTGAGTGGAGCGCGCGCAATGTTTTTTTTTTTTTTTTTTT
 T 3'); double-stranded cDNA was ligated
 adaptors (Pharmacia), digested with Not
 the Not I and Eco RI sites of the modifi
 RNA provided by Dr. Minoru Ko, Wayne Sta
 constructed and normalized by Bento Soares
 Bonaldo."

BASE COUNT

103 a 101 c 98 g 96 t

ORIGIN

Query Match 1.54; Score 23; DB 9; Length 398;
 Best Local Similarity 100.0%; Pred No. 3.1;
 Matches 23; Conservative 0; Mismatches 0; Indels 0;

QY 1076 GGCTACCTCTGTGACCGAGCTT 1098

Db 147 GGCTACCTCTGTGACCGAGCTT 169

RESULT 34

LOCUS

BB583702 655 bp mRNA linear
 BB583702 RIKEN full-length enriched, adult male

CDNA clone 9130603.18 5', mRNA sequence.
 BB583702
 BB583702.2 GI:16449603

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2001)
 On Nov 30, 2000 this sequence version replaced 1
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research
 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@scf.riken.go.jp/
 URL: http://genome.scf.riken.go.jp/
 Carninci, P., Shibata, Y., Hayashizaki, N., Sugahara, Y.,
 M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap trapped so
 prepare full-length cDNA libraries for rapid dis
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujisaki, S., Inoue, K., Toyama, Y., Izawa,
 Matsubara, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Ii,
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, J.

Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multiplexedly sequenced genome
10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia. Real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamazawa, T., Aizawa,
F., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome-gsc.riken.go.jp/>) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

FEATURES

Source

Location/Qualifiers
1..655
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="q1000418"
/clone_11b="RIKEN full length cDNA, adult male, testis"
/sex="male"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site 1: SalI, site 2: BamHI. cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
project of Genome Exploration Research group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer (5'
CAAGAGACAGACATTCAGACATCTTTTCTTTTCTTTCVN 3'). cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 185.0. Second
strand cDNA was prepared with the primer adapter of
sequence (5' GAGAGACATTCGACATTAATTAATTCGCCCCCCCC
3'). cDNA was cloned into the XhoI and BamHI sites.
Vector: a modified pRisc(+) after bulk excision
from Lambda E1c I. Cloning sites, 5' end: SalI; 3' end:
BamHI"

BASE COUNT 176 a 160 c 165 g 153 t 1 others

Query Match: 1.5%, Score 23, DB 1c, Length 655,
Best Local Similarity 100.0%, Pred. No. 2.9;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1476 GATTCACCTCTCTGACCGACCTT 1498
|||||
Db 613 GGTTCACCTCTCTGACCGACCTT 635

RESULT 35
A1208287 417 bp mRNA linear EST 30-NOV-1998
LAVUS
DEFINITION 995911.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1839524
ACCESSION A1208287
VERSION A1208287.1 GI:3770229
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
ADDITIONAL
TITLE
JOURNAL
COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 417)
NCI Code <http://www.ncbi.nlm.nih.gov/ncicomp>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: raeb@rmail.nih.gov
cDNA library preparation: M. Bento Soares, Ph.D., M. Fatima Ronaldi
Ph.D.

cDNA library Arrayed by: Eric Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the 1 M A G E Consortium/1MAG at
www.tpo.1m1.gov/hfpg/umag/umag.html
Insert length: 523 Std Error: 0.00
Seq primer: 400P from Glibco.

FEATURES

Source

Location/Qualifiers
1..417
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1839524"
/clone_11b="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pTZ19-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: EcoRI. 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I oligo(dT) primer (5'
TCTTACCAATTTGAGAGAGAGAGAGAGAGAGATTTTCTTTTCTTTCVN 3').
Double-stranded cDNA was ligated to EcoRI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and EcoRI sites of the modified pTZ19 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Ronaldi.

BASE COUNT 113 a 77 c 104 g 123 t

Query Match: 1.4%, Score 22, DB 9; Length 417;
Best Local Similarity 100.0%, Pred. No. 9.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1477 GATTCGCTTTCTGACCTGTT 1498
|||||
Db 394 GATTCGCTTTCTGACCTGTT 373

RESULT 36
A0433394/c 555 bp DNA linear GSS 31-MAR-1999
LAVUS
DEFINITION HS_5050_H1_F08_77A RGC1-11 Human Male BAC Library Homo sapiens
genomic clone Plate 626 Col=15 Row=L. DNA sequence.
ACCESSION A0433394
VERSION A0433394.1 GI:4543729
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 555)
AUTHORS Mahairas, G. G., Wallace, J. C., Smith, K., Swartz, L. S., Holzman, T.,
Keller, A., Shaker, P., Furlong, J., Young, J., Zhao, S., Adams, M. D., and
Hood, L.
TITLE sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc Natl Acad Sci U S A 96 (17), 9739-9744 (1999)
PMID 99380589
COMMENT Contact: Mahairas G., Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887

BASE COUNT 308 a 226 c 247 g 287 t 3 others
 ORIGIN

Query Match 1.4% Score 22; DH 13; Length 1071;
 Best Local Similarity 100.0%; Pred. No. 8.7;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1477 CATTGGCTGTTCTGAGTCT 1498
 |||||||
 Db 116 CATTGGCTGTTCTGAGTCT 137

RESULT 42
 A1602007/c 210 bp mRNA linear EST 21-APR-1999
 LOCUS

DEFINITION U1-R-Y0-uv-g-02-0-01 s2 U1-R-Y0 Rattus norvegicus cDNA clone
 A1602007
 VERSION A1602007
 KEYWORDS EST.

SOURCE Norway rat.
 ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 210)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.iowa.edu

The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 clones will be available through Research Genetics (www.resgen.com)
 Seq primer: M13 forward.
 Location/Qualifiers

1..210
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="U1-R-Y0-uv-g-02-0-01"
 /clone_lib="U1-R-Y0"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73b-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The U1-R-Y0
 library is a subtracted library derived from an
 individually-tagged normalized whole-eye (minus the lens)
 library. The driver for the subtraction consisted of a
 pool of all previous libraries (U1-R-A0, U1-R-A1, U1-R-E0,
 U1-R-E1, U1-R-C0, and U1-R-C1). The tag is a string of
 3-5 nucleotides present between the Not I site and the
 oligo-dT track which allows identification of the library
 of origin of a clone within the mixture. The subtracted
 library (U1-R-Y0) was constructed as follows: PCR
 amplified cDNA inserts from previous library clones from
 which 3' ESTs had been derived were used as a driver in a
 hybridization with the normalized whole-eye library in
 the form of single-stranded circles. The remaining
 single-stranded circles (subtracted library) was purified
 by hydroxyapatite column chromatography, converted to
 double-stranded circles and electroporated into DH10B
 bacteria (Life Technologies) to generate the U1-R-Y0

library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Res.
 1996)."
 BASE COUNT 55 a 41 c 44 g 70 t
 ORIGIN

Query Match 1.4% Score 21; DH 9; Length 21
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 21; Conservative 0; Mismatches 0; Indels

CY 1522 TGTGTACTGCAAAAAA 1542
 |||||||
 Db 29 TGTGTACTGCAAAAAA 9

RESULT 43
 A0270139/c 477 bp DNA linear
 LOCUS

DEFINITION HS-2060_A1_H12_T7 c11 Approved Human Genomic Spee
 sapiens genomic clone Plate-2060 Col-23 Row-0, DNT
 A0270139
 VERSION A0270139.1 GI:3824204
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 377)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Spartz, J.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, J.,
 Hood, L.

Sequence-tagged connectors: A sequence approach to
 scanning the human genome
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744
 99380589
 COMMENT Contact: Mahairas G.G., Wallace J.C., Hood L.
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, US
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2060 row: 0 column: 23
 Class: BAC ends
 High quality sequence stop: 477.
 Location/Qualifiers

1..377
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Plate 2060 Col-23 Row-0"
 /clone_lib="c11 Approved Human Genomic Sp
 /sex="male"
 /note="Organ: Sperm; Vector: pHELOBAC11;
 E-Coli DH10B"
 BASE COUNT 99 a 102 c 91 t 3 others
 ORIGIN

Query Match 1.4% Score 21; DH 17; Length 47
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 21; Conservative 0; Mismatches 0; Indels

UY 1406 CCACAGACGCGATGAGCGTC 1426
 |||||||
 Db 236 CCACAGACGCGATGAGCGTC 216

RESULT 44
 A1764709/c 404 bp mRNA linear
 LOCUS
 DEFINITION U1-R-Y0-ace-b-07-0-01 s1 U1-R-Y0 Rattus norvegicus
 U1-R-Y0-ace-b-07-0-01 s1, mRNA sequence.
 ACCESSION A1764709

VERSION A1764709.1 GI:5210644
 KEYWORDS EST
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 404)
 AUTHORS Ronalds, M.F., Lennon, G., and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MR
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the
 normalized eye library cDNA library preparation; M.B. Soares Lab
 Clone distribution: clones will be available through Research
 Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLYA-yes.

FEATURES
 source
 1 404
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="U1-R-Y0-acc-b-07-0-01"
 /clone_lib="U1-R-Y0"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pTZ19-Pac (Pharmacia) with a modified
 polylinker site. Not I Site 2. Eco RI: The U1-R-Y0
 library is a subtracted library derived from an
 individually-tagged normalized whole-eye (minus the lens)
 library. The driver for the subtraction consisted of a
 pool of all previous libraries (U1-R-A0, U1-R-A1, U1-R-E0,
 U1-R-E1, U1-R-E0, and U1-R-G1). The tag is a string of
 3-5 nucleotides present between the Not I site and the
 oligo-dT track which allows identification of the library
 of origin of a clone within the mixture. The subtracted
 library (U1-R-Y0) was constructed as follows: pTZ
 amplified cDNA inserts from previous library clones for
 which 3' ESTs had been derived were used as a driver in a
 hybridization with the normalized whole eye library in
 the form of single-stranded circles. The remaining
 single-stranded circles (subtracted library) was purified
 by hydroxyapatite column chromatography, converted to
 double-stranded circles and electroporated into DH10B
 bacteria (Life Technologies) to generate the U1-R-Y0
 library. This procedure has been previously described
 (Ronalds, Lennon and Soares, Genome Research 6: 791-806,
 1996)
 TAG_LIB=U1-R-Y0
 TAG_TISSUE=Eye
 TAG_SEQ=CATTG"

BASE COUNT 93 a 105 c 95 g 111 t
 ORIGIN

Query Match 1.4%; Score 21; DB 9; Length 404;
 Best Local Similarity 100.0%; Prod. No. 31;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 1522 TCTGTACGCGCAAAAAA 1542
 |||

pb 29 TCTGTACGCGCAAAAAA 9

RESULT 45
 AC663408 431 bp DNA linear GSS 23-JUN-1999
 LOCUS Hs_2164_H21609_275_01T Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone plate 2160 Col-12 Row-H, DNA sequence.
 ACCESSION AC663408
 VERSION AC663408.1 GI:5171176
 KEYWORDS GSS
 SOURCE human:
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 431)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Kolter, A., Stokoe, R., Furlong, J., Young, T., Zhao, S., Adams, M.D., and
 Hood, L.
 TITLE Sequenced-tagged cDNAs: A sequence approach to mapping and
 screening the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com),
 BAC end Web Server: <http://www.bac.washington.edu>
 Plate: 2160 Row: H Column: 12
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 431

FEATURES
 source
 1 431
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="H21609-275_001_12 Row-H"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /sex="male"
 /note="Organ: sperm; vector: pTZ19AC11; BAC clones in
 E-Coli DH10B"

BASE COUNT 124 a 101 c 116 g 1 others
 ORIGIN

Query Match 1.4%; Score 21; DB 17; Length 431;
 Best Local Similarity 100.0%; Prod. No. 31;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Caps 0;
 1477 CATTGCTGTCTGCACTTC 1497
 |||

db 227 CATTGCTGTCTGCACTTC 247

Search completed: December 7, 2002, 11:02:15
 Job time: 2098 secs

Q7 1472 TCTTCATTGGCTGTTCTGA 1493
|||||
Db 64 TCTTCATTGGCTGTTCTGA 42

RESULT 7
US-09-604-287A-29/C
Sequence 29, Application US/09604287A
Patent No. US20020054872A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.47067
CURRENT FILING DATE: 2000-06-22
NUMBER OF SEQ ID NOS: 489
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 29
LENGTH: 301
TYPE: DNA
ORGANISM: Homo sapien
US-09-604-287A-29

Query Match 1.28; Score 19; DB 10; Length 301;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 1060 GTGCATATGGCTGCACGGC 1078
|||||
Db 265 GTGCATATGGCTGCACGGC 247

RESULT 8
US-09-339-438-29/C
Sequence 29, Application US/09339338A
Patent No. US20020102602A1
GENERAL INFORMATION:
APPLICANT: Yugu, Jiang
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THE USE
FILE REFERENCE: 210121.47067
CURRENT APPLICATION NUMBER: US/09/438,438A
CURRENT FILING DATE: 1999-05-23
NUMBER OF SEQ ID NOS: 315
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 29
LENGTH: 301
TYPE: DNA
ORGANISM: Homo sapien
US-09-339-438-29

Query Match 1.28; Score 19; DB 10; Length 301;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 1060 GTGCATATGGCTGCACGGC 1078
|||||
Db 265 GTGCATATGGCTGCACGGC 247

RESULT 9
US-10-007-805-29/C
Sequence 29, Application US/10007805

Patent No. US20020150581A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Fung, Gary K.
APPLICANT: Vedvick, Thomas S.
APPLICANT: McNeill, Patricia D.
APPLICANT: Durham, Margarita
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470610
CURRENT APPLICATION NUMBER: US/10/007,805
CURRENT FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 593
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 29
LENGTH: 301
TYPE: DNA
ORGANISM: Homo sapien
US-10-007-805-29

Query Match 1.28; Score 19; DB 12; Length 301;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q7 1060 GTGCATATGGCTGCACGGC 1078
|||||
Db 245 GTGCATATGGCTGCACGGC 247

RESULT 10
US 09 864 761 1342
Sequence 1342, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon C.
APPLICANT: Kunk, David R.
APPLICANT: Hanzel, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Neomix-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 90/190,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/7246,459
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00659
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00655
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

```

PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00679
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine v0rs. 1.1
SEQ ID NO: 13242
LENGTH: 520
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC017079.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SYCNAL - 1
US-09-854-761-13242

```

```

Query Match          12%: Score 19, PH 10, Length 520;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 474 AACGTGAGTGGTGAGCT 492
    ||||||||||||||||
DB 169 AACGTGAGTGGTGAGCT 187

```

RESULT 11

```

US-09-905-291A-376
Sequence 376, Application US/09905291A
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottlieb, Mary E.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavitt, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William L.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905,291A
PRIOR FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26

```

```

PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-14
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/209214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28413
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO: 376
LENGTH: 997
TYPE: DNA
ORGANISM: Homo Sapien
US-09-905-291A-376

```

```

QY 1060 GTGGCAATGGTGAGCGC 1078
    ||||||||||||||||
DB 552 GTGGCAATGGTGAGCGC 570

```

RESULT 12

```

US-09-909-320-376
Sequence 376, Application US/09909320
Patent No. US2002013240A1
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gottlieb, Mary E.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavitt, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey

```

APPLICANT: Wood, William, L.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21095
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/22084
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28413
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/29599
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO: 376
LENGTH: 997
TYPE: DNA
ORGANISM: Homo Sapien
US-09-909-320-376

Query Match: 1 26; Score 13; DB 10; Length 997;
Best Local Similarity 100.0%; Pred No. 7.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1060 GTGGCAATGGTGACGCGC 1078
|||||
Db 552 GTGGCAATGGTGACGCGC 570

RESULT 13
US-09-909-088B-376
Sequence 376, Application US/09909088B
Patent No. US20020146709A1
GENERAL INFORMATION:
APPLICANT: Clementech, Inc
APPLICANT: Ashkenazi, Avi
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlisen, Mary F.
APPLICANT: Goddard, A.

APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, Jr.
APPLICANT: Kistwin, Ivar J.
APPLICANT: Mather, Jennie B.
APPLICANT: Pan, James
APPLICANT: Paoletti, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Thomas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, L.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21095
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28413
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO: 376
LENGTH: 997
TYPE: DNA
ORGANISM: Homo Sapien
US-09-909-088B-376

Query Match: 1 26; Score 13; DB 10; Length 997;
Best Local Similarity 100.0%; Pred No. 7.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1060 GTGGCAATGGTGACGCGC 1078
|||||
Db 552 GTGGCAATGGTGACGCGC 570

RESULT 14
US-10-052-586-23
Sequence 23, Application US/10052586
Patent No. US20020127584A1
GENERAL INFORMATION:
APPLICANT: Goddard, A.

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C1
CURRENT APPLICATION NUMBER: US/10/052,586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063564
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063734
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066456
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069345
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077652
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078846
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078849
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107

PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080343
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081848
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083422
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083579
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086302
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086466
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087779
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04

```

1  CUREP: APL1-ALL-130868: 1091-05-23
2  CUREP FILING DATE: 2001-05-23
3  PRIOR APPLICATION NUMBER: US 69/180,312
4  PRIOR FILING DATE: 2000-02-04
5  PRIOR APL1-ALL-130868: 1091-05-26
6  PRIOR FILING DATE: 2000-05-26
7  PRIOR APPLICATION NUMBER: US 09/632,366
8  PRIOR FILING DATE: 2000-06-03
9  PRIOR APPLICATION NUMBER: GB 24263,6
10 PRIOR FILING DATE: 2000-10-04
11 PRIOR APPLICATION NUMBER: US 60/736,354
12 PRIOR FILING DATE: 2000-09-27
13 PRIOR APPLICATION NUMBER: PCT/US01/00666
14 PRIOR FILING DATE: 2001-01-30
15 PRIOR APPLICATION NUMBER: PCT/US01/00667
16 PRIOR FILING DATE: 2001-01-30
17 PRIOR APPLICATION NUMBER: PCT/US01/00664
18 PRIOR FILING DATE: 2001-01-30
19 PRIOR APPLICATION NUMBER: PCT/US01/00669
20 PRIOR FILING DATE: 2001-01-30
21 PRIOR APPLICATION NUMBER: PCT/US01/00665
22 PRIOR FILING DATE: 2001-01-30
23 PRIOR APPLICATION NUMBER: PCT/US01/00668
24 PRIOR FILING DATE: 2001-01-30
25 PRIOR APPLICATION NUMBER: PCT/US01/00663
26 PRIOR FILING DATE: 2001-01-30
27 PRIOR APPLICATION NUMBER: PCT/US01/00662
28 PRIOR FILING DATE: 2001-01-30
29 PRIOR APPLICATION NUMBER: PCT/US01/00661
30 PRIOR FILING DATE: 2001-01-30
31 PRIOR APPLICATION NUMBER: PCT/US01/00670
32 PRIOR FILING DATE: 2001-01-30
33 PRIOR APPLICATION NUMBER: US 72/234,687
34 PRIOR FILING DATE: 2000-09-21
35 PRIOR APPLICATION NUMBER: US 09/604,498
36 PRIOR FILING DATE: 2000-06-30
37 PRIOR APPLICATION NUMBER: US 09/774,203
38 PRIOR FILING DATE: 2001-01-29
39 NUMBER OF SEQ ID NOS: 49117
40 SOFTWARE: Autonomax Sequence Listing Engine vers. 1.1
41 SEQ ID NO: 4689
42 LENGTH: 423
43 TYPE: DNA
44 ORGANISM: Homo sapiens
45 FEATURE:
46 OTHER INFORMATION: MAP TO AL109798.13
47 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.2
48 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
49 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.9
50 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.1
51 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
52 OTHER INFORMATION: EXPRESSED IN BT774, SIGNAL = 6
53 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.5
54 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.1
55 OTHER INFORMATION: EXPRESSED IN HPLA, SIGNAL = 5.3
56 OTHER INFORMATION: EXPRESSED IN HH109, SIGNAL = 5.6
57 US-09-864-761-3689
58
59 Query Match 1.28, Score 14, DB 10, Length 423;
60 Best Local Similarity 100.0%, Fred. No. 23;
61 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0
62
63 Q7 1209 TCTGTAAAAGCGCCGAG 1256
64 |||||||
65 DB 178 TCTGTAAAAGCGCCGAG 161
66
67 RESULT 16
68 US-09-954-456-1645/C
69 Sequence 1645, Application US/09954456
70 Patent No. US20020115057A1
71 APPLICANT: Young, Paul

```



```

: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cando
: TITLE OF INVENTION: Sels
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954,456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,863
: PRIOR FILING DATE: 2000-09-27
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO: 1685
: LENGTH: 456
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-954-456-1685

Query Match 1.28; Score 18; DB 10; Length 456.
Best Local Similarity 100.0%; Pred. No. 23;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 503 TCCAGCTGCAGGAGATC 520
|||||
DB 81 TCCAGCTGCAGGAGATC 64

RESULT 17
US-10-044-090-690/c
: Sequence 690, Application US/10044090
: Patent No. US20020137081A1
: GENERAL INFORMATION:
: APPLICANT: Olga Bandman
: TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
: FILE REFERENCE: PA-002A US
: CURRENT APPLICATION NUMBER: US/10/044,090
: CURRENT FILING DATE: 2002-01-09
: NUMBER OF SEQ ID NOS: 850
: SOFTWARE: PERL Program
: SEQ ID NO: 690
: LENGTH: 1389
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incycle ID No. US20020137081A1 480287.8
US-10-044-090-690

Query Match 1.28; Score 18; DB 12; Length 1389.
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 TTCGTGTTGTTGCTG 175
|||||
DB 964 TTCGTGTTGTTGCTG 947

RESULT 18
```

```

US-09-729-674-19
: Sequence 19, Application US/09729674
: Patent No. US20010039335A1
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: McCoy, John M.
: APPLICANT: Lavallee, Edward E.
: APPLICANT: Collins-Ragley, Lisa A.
: APPLICANT: Evans, Cheryl
: APPLICANT: Mebery, David
: APPLICANT: Treacy, Maurice
: APPLICANT: Apostoluo, Michael J.
: APPLICANT: Steuninger II, Robert J.
: APPLICANT: Spaulding, Virki
: APPLICANT: Wong, Gordon G.
: APPLICANT: Clark, Hilary
: APPLICANT: Fectel, Kim
: APPLICANT: Genetics Institute, Inc.
: TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES
: FILE REFERENCE: 6055-64X
: CURRENT APPLICATION NUMBER: US/09/729,674
: CURRENT FILING DATE: 2000-12-04
: PRIOR APPLICATION NUMBER: US/759,330
: PRIOR FILING DATE: 2000-03-30
: NUMBER OF SEQ ID NOS: 283
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 19
: LENGTH: 1656
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-729-674-19

Query Match 1.28; Score 18; DB 10; Length 1656.
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1525 TGTACTGCACAAAAAAA 1542
|||||
DB 1596 TGTACTGCACAAAAAAA 1613

RESULT 19
US-09-764-869-2142/c
: Sequence 2142, Application US/09764869
: Patent No. US20020061521A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: P0007
: CURRENT APPLICATION NUMBER: US/09/764,869
: PRIOR APPLICATION DATA REMOVED - refer to PAIN or file with
: NUMBER OF SEQ ID NOS: 2442
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO: 2142
: LENGTH: 6800
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-764-869-2142

Query Match 1.28; Score 18; DB 10; Length 6800.
Best Local Similarity 100.0%; Pred. No. 24;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 991 ACAGCTGCAGACTTCAG 1008
|||||
DB 1443 ACAGCTGCAGACTTCAG 1426

RESULT 20
US-09-917-800A-502
: Sequence 502, Application US/09917800A
: Patent No. US20020119462A1
```

```

: GENERAL INFORMATION:
: APPLICANT: Mendrick, Donna
: APPLICANT: Porter, Mark
: APPLICANT: Johnson, Kory
: APPLICANT: Castle, Arthur
: APPLICANT: Elashoff, Michael
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Molecular Toxicology Modeling
: FILE REFERENCE: 44921-5038-US
: CURRENT APPLICATION NUMBER: US/09/917,800A
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/222,040
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: US 60/222,880
: PRIOR FILING DATE: 2000-11-02
: PRIOR APPLICATION NUMBER: US 60/240,029
: PRIOR FILING DATE: 2001-05-11
: PRIOR APPLICATION NUMBER: US 60/250,645
: PRIOR FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: US 60/292,336
: PRIOR FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: US 60/295,798
: PRIOR FILING DATE: 2001-06-06
: PRIOR APPLICATION NUMBER: US 60/297,457
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,984
: PRIOR FILING DATE: 2001-06-19
: PRIOR APPLICATION NUMBER: US 60/303,459
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 1740
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 502
: LENGTH: 7420
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURES:
: OTHER INFORMATION: Genbank Accession No. US20020114462A1 Approved
US-09-917-800A-502

Query Match          1.2%, Score 18, DB 10, Length 7420
Best Local Similarity 100.0%, Pred. No. 26;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 450 TCTACCAAGACCTGACCC 467
      |||||||
DB 7286 TCTACCAAGACCTGACCC 7303

RESULT 21
US-09-954-456-2006
: Sequence 2006, Application US/09/954,456
: Patent No. US20020115057A1
: GENERAL INFORMATION:
: APPLICANT: Young, Paul
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-75
: CURRENT APPLICATION NUMBER: US/09/954,456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,648
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27

```

```

: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,863
: PRIOR FILING DATE: 2000-09-27
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: Patent In version 4.0
: SEQ ID NO 2006
: LENGTH: 7787
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-954-455-2006

Query Match          1.2%, Score 18, DB 10, Length 7787;
Best Local Similarity 100.0%, Pred. No. 26;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 450 TCTACCAAGACCTGACCC 467
      |||||||
DB 7388 TCTACCAAGACCTGACCC 7405

RESULT 22
US-09-962-832-154
: Sequence 154, Application US/09/962,832
: Patent No. US20020110821A1
: GENERAL INFORMATION:
: APPLICANT: Ebner, Reinhard
: TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-74
: CURRENT APPLICATION NUMBER: US/09/962,832
: CURRENT FILING DATE: 2001-09-25
: PRIOR APPLICATION NUMBER: US/60/235,077
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,290
: PRIOR FILING DATE: 2000-09-25
: NUMBER OF SEQ ID NOS: 259
: SOFTWARE: Patent In version 4.0
: SEQ ID NO 154
: LENGTH: 302250
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-962-832-154

Query Match          1.2%, Score 18, DB 10, Length 302250;
Best Local Similarity 100.0%, Pred. No. 31;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1316 CCTTTTCAATGCTGCTGACT 1333
      |||||||
DB 160218 CCTTTTCAATGCTGCTGACT 160235

RESULT 23
US-09-560-863-384
: Sequence 384, Application US/09/560,863
: Patent No. US20020110809A1
: GENERAL INFORMATION:
: APPLICANT: Noblis, Michael C.
: APPLICANT: Gabbrowicz, Brian
: APPLICANT: Sands, Arthur T.
: TITLE OF INVENTION: No. US20020110809A1; Human Polynucleotides and the
: TITLE OF INVENTION: Polynucleotides Encoded Thereby
: FILE REFERENCE: LEX-0018-USA
: CURRENT APPLICATION NUMBER: US/09/560,863
: CURRENT FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/132,408
: PRIOR FILING DATE: 1999-04-30
: NUMBER OF SEQ ID NOS: 1008
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 384

```

```

: LENGTH: 353
: TYPE: DNA
: ORGANISM: homo sapiens
: US-09-560-863-384

Query Match
Best Local Similarity 100.0%; Score 17; DB 10; Length 353;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1526 GTAACTGCAAAAAAAA 1542
      |||
Db 337 GTAACTGCAAAAAAAA 353

RESULT 24
US-09-864-761-2789/c
: Sequence 2789, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aeonica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 2789
: LENGTH: 381
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC009229.1
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3

```

```

: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
: OTHER INFORMATION: EXPRESSED IN H1474, SIGNAL = 2.9
: OTHER INFORMATION: EXPRESSED IN HONE MARROW, SIGNAL = 4.0
: OTHER INFORMATION: EXPRESSED IN H9100, SIGNAL = 2.8
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.0
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.5
: OTHER INFORMATION: EXPRESSED IN HEK293, SIGNAL = 3.6
: US-09-864-761-2789

Query Match
Best Local Similarity 100.0%; Score 17; DB 10; Length
Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 382 GGCATCTTGTCTCAAG 394
      |||
Db 348 GGCATCTTGTCTCAAG 342

RESULT 25
US-09-783-590-7957
: Sequence 7957, Application US/09783590
: Patent No. US20020110850A1
: GENERAL INFORMATION:
: APPLICANT: Dillon, Patrick J.
: APPLICANT: Haseltine, William A.
: APPLICANT: Li, Haodong
: APPLICANT: Rosen, Craig A.
: APPLICANT: Ruben, Steven M.
: TITLE OF INVENTION: Human Genes, Sequences, and Expression
: FILE REFERENCE: PO-16,201
: CURRENT APPLICATION NUMBER: US/09/783,590
: CURRENT FILING DATE: 2000-02-15
: PRIOR APPLICATION NUMBER: 08/420,856
: PRIOR FILING DATE: 1995-04-12
: PRIOR APPLICATION NUMBER: 08/446,741
: PRIOR FILING DATE: 1994-11-21
: NUMBER OF SEQ ID NOS: 12485
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 7957
: LENGTH: 442
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (10)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (26)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (189)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (203)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (286)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (313)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (316)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (327)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature
: LOCATION: (338)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: misc feature

```

```

: LOCATION: (394)
: OTHER INFORMATION: n equals a,t,q, or c
: NAME/KEY: misc feature
: LOCATION: (395)
: OTHER INFORMATION: n equals a,t,q, or c
: NAME/KEY: misc feature
: LOCATION: (398)
: OTHER INFORMATION: n equals a,t,q, or c
: NAME/KEY: misc feature
: LOCATION: (404)
: OTHER INFORMATION: n equals a,t,q, or c
: NAME/KEY: misc feature
: LOCATION: (427)
: OTHER INFORMATION: n equals a,t,q, or c
: NAME/KEY: misc feature
: LOCATION: (435)
: OTHER INFORMATION: n equals a,t,q, or c
US-09-783-590-7957

```

```

Query Match: 1.1%; Score 17; DB 10; Length 412;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 866 TTCTTCCAGGAACGTG 882
    11111111111111111111
DB 291 TTCTTCCAGGAACGTG 307

```

```

RESULT 26
US-09-864-761-1416/c
: Sequence 1416, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David P.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DEPENDED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aecmiga-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/432,366
: PRIOR FILING DATE: 2000-04-03
: PRIOR APPLICATION NUMBER: GR 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,259
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687

```

```

: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 1416
: LENGTH: 452
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO APO2082.1
: OTHER INFORMATION: EXPRESSED IN HELLO, SIGNAL = 2.6
: OTHER INFORMATION: EXPRESSED IN HEPL, SIGNAL = 3.3
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.4
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.6
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.2
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
US-09-864-761-1416

```

```

Query Match: 1.1%; Score 17; DB 10; Length 452;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1079 TCACCTTGTCCACCCAG 1095
    11111111111111111111
DB 242 TCACCTTGTCCACCCAG 226

```

```

RESULT 27
US-09-974-300-2991
: Sequence 2991, Application US/09974300
: Patent No. US20020146721A1
: GENERAL INFORMATION:
: APPLICANT: Berkas, Randy M.
: APPLICANT: Clausen, ID Groth
: TITLE OF INVENTION: Methods For Monitoring Multiple Gene
: FILE REFERENCE: 10085,500-US
: CURRENT APPLICATION NUMBER: US/09/974,300
: CURRENT FILING DATE: 2001-10-05
: PRIOR APPLICATION NUMBER: 09/490,538
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 69/279,526
: PRIOR FILING DATE: 2001-03-27
: NUMBER OF SEQ ID NOS: 8481
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2991
: LENGTH: 573
: TYPE: DNA
: ORGANISM: Bacillus licheniformis
US-09-974-300-2991

```

```

Query Match: 1.1%; Score 17; DB 10; Length 573;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 591 ACGACGACGTGACGCGG 607
    11111111111111111111
DB 245 ACGACGACGTGACGCGG 252

```

```

RESULT 28
US-09-864-761-17795/c
: Sequence 17795, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.

```

```

APPLICANT: Hank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecomic-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GR 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 17795
LENGTH: 577
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC009229.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.6
OTHER INFORMATION: SWISSPROT HIT: Q16674, EVALUATE 1.00e-111
OTHER INFORMATION: NT HIT: Q111440100, EVALUATE 0.00e+00
OTHER INFORMATION: EST_HUMAN HIT: A0120297.1, EVALUATE 0.00e+00
US-09-864-761-17795

```

```

Query Match 1.18, Score 17, DH 10, Length 577;
Best Local Similarity 100.08; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
382 GAGCATCTTGTCCAG 398
|||||

```

```

Db 308 GAGCATCTTGTCCAG 292
RESULT 29
US-09-864-761-16207
Sequence 16207, Application 05/0984761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC
FILE REFERENCE: Aecomic-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GR 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 16207
LENGTH: 592
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL145905.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
US-09-864-761-16207

```

```

Query Match 1.18, Score 17, DH 10, Length 592;
Best Local Similarity 100.08; Pred. No. 75;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
358 CATGCTTGTGGG 374
|||||
Db 269 CATGCTTGTGGG 285

```

RESULT 30

```

US-09-764-846-76
: Sequence 76, Application US/03764846
: Patent No. US20020102638A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, proteins, and Antibodies
: FILE REFERENCE: PT212
: CURRENT APPLICATION NUMBER: US/09/764,846
: PRIOR FILING DATE: 2001-01-17
: Prior application data removed - consult TADM or file wrapper
: NUMBER OF SEQ ID NOS: 348
: SOFTWARE: Patent Ver. 2.0
: SEQ ID NO 76
: LENGTH: 759
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (28)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (30)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (663)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (690)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (700)
: OTHER INFORMATION: n equals a,t,g, or c
: NAME/KEY: SITE
: LOCATION: (711)
: OTHER INFORMATION: n equals a,t,g, or c
US-09-764-846-76

```

Query Match

```

1.1%, Score 17, DB 10, Length 759;
Best local Similarity 100.0%, Pred. No. 76;
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

```

```

QY 975 AGGAGCAACTCTCTCA 991
    |||||
Db 217 AGGAGCAACTCTCTCA 233

```

RESULT 31

```

US-09-861-893-48/7
: Sequence 39, Application US/09861893
: Patent No. US20020045257A1
: GENERAL INFORMATION:
: APPLICANT: Feinberg, Andrew
: APPLICANT: Strichman, Almasanu, Liola
: APPLICANT: Jiang, Shao
: TITLE OF INVENTION: METHODS FOR ASSAYING GENE IMPRINTING AND
: FILE REFERENCE: 01107, 00128
: CURRENT APPLICATION NUMBER: US/09/861,893
: PRIOR FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: 60/206,158
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: 60/206,161
: PRIOR FILING DATE: 2000-05-22
: NUMBER OF SEQ ID NOS: 77
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 38
: LENGTH: 788
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-861-893-48

```

Query Match

```

1.1%, Score 17, DB 10, Length 788;

```

```

Best local Similarity 100.0%, Pred. No. 76;
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

```

```

QY 293 GTGCAGTACCTGGGCT 309
    |||||
Db 742 GTGCAGTACCTGGGCT 726

```

RESULT 32

```

US-09-916-790-3
: Sequence 3, Application US/09916790
: Patent No. US20020661573A1
: GENERAL INFORMATION:
: APPLICANT: Meyers, Rachael
: APPLICANT: Kapellor-Libermann, Rosana
: APPLICANT: Sittos Santiago, Immaculada
: TITLE OF INVENTION: 18431 AND 32374, NOVEL HUMAN PROTEIN
: FILE REFERENCE: 38152002700
: CURRENT APPLICATION NUMBER: US/09/916,790
: PRIOR FILING DATE: 2001-07-27
: PRIOR APPLICATION NUMBER: 60/221,543
: PRIOR FILING DATE: 2000-07-28
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 1041
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-916-790-3

```

Query Match

```

1.1%, Score 17, DB 10, Length 1041;
Best local Similarity 100.0%, Pred. No. 77;
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

```

```

QY 293 GTGCAGTACCTGGGCT 309
    |||||
Db 229 GTGCAGTACCTGGGCT 245

```

RESULT 33

```

US-09-938-842A-114
: Sequence 134, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Krebs, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: FILE REFERENCE: SCRIPT300-3
: CURRENT APPLICATION NUMBER: US/09/938,842A
: PRIOR FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 09/421,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 09/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 134
: LENGTH: 1059
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-114

```

Query Match

```

1.1%, Score 17, DB 9, Length 1059;
Best local Similarity 100.0%, Pred. No. 78;
Matches 17, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

```

```

QY 1396 CCAATGCAATCCAGACA 1412
    |||||

```

DB 828 CCAATGCGATCGAGACA 844

RESULT 34

```
US-09-833-381-71
: Sequence 71, Application US/09833381
: Patent No. US20020132090A1
: GENERAL INFORMATION:
: APPLICANT: Robison, Keith E.
: TITLE OF INVENTION: No. US20020132090A1: Nucleic Acid and Protein Homologs
: FILE REFERENCE: 5800-119
: CURRENT APPLICATION NUMBER: US/09/833,381
: PRIOR FILING DATE: 2001-04-11
: PRIOR APPLICATION NUMBER: 09/516,448
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 2050
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 71
: LENGTH: 1154
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1) (1154)
: OTHER INFORMATION: n - A,T,C or G
US-09-833-381-71
```

Query Match

```
1.18: Score 17; DB 10; Length 1154,
Best Local Similarity 100.0%; Pred. No. 78;
Matches 17; Conservative 0; Mismatches 0;
```

QY 243 TGAACATCTGGAGGAG 259

DB 269 TGAACATCTGGAGGAG 285

RESULT 45

```
US-09-917-800A-1377
: Sequence 1377, Application US/09917800A
: Patent No. US2002019462A1
: GENERAL INFORMATION:
: APPLICANT: Mendrick, Donna
: APPLICANT: Porter, Mark
: APPLICANT: Johnson, Kory
: APPLICANT: Castle, Arthur
: APPLICANT: Elashoff, Michael
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Molecular Toxicology Modeling
: FILE REFERENCE: 44921-5038-05
: CURRENT APPLICATION NUMBER: US/09/917,800A
: PRIOR FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/222,040
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: US 60/222,880
: PRIOR FILING DATE: 2000-11-02
: PRIOR APPLICATION NUMBER: US 60/290,029
: PRIOR FILING DATE: 2001-05-11
: PRIOR APPLICATION NUMBER: US 60/290,645
: PRIOR FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: US 60/292,336
: PRIOR FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: US 60/295,798
: PRIOR FILING DATE: 2001-06-06
: PRIOR APPLICATION NUMBER: US 60/297,457
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,884
: PRIOR FILING DATE: 2001-06-19
: PRIOR APPLICATION NUMBER: US 60/303,459
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 1740
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1377
: LENGTH: 1164
```

```
: TYPE: DNA
: ORGANISM: Rattus norvegicus
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020119462A1
US-09-917-800A-1377
```

Query Match

```
1.18: Score 17; DB 10; Length 1164,
Best Local Similarity 100.0%; Pred. No. 78;
Matches 17; Conservative 0; Mismatches 0;
```

QY 1354 CTTCGAGCTGCGAGCA 1370

DB 992 CTTCGAGCTGCGAGCA 1008

RESULT 36

```
US-10-100-745-3
: Sequence 3, Application US/10100745
: Patent No. US20020115169A1
: GENERAL INFORMATION:
: APPLICANT: Olson, Mark
: APPLICANT: Szebenl, Atilla
: TITLE OF INVENTION: METHODS FOR STABILIZATION AND
: FILE REFERENCE: 13121.000203
: CURRENT APPLICATION NUMBER: US/10/100,745
: PRIOR FILING DATE: 2002-04-18
: PRIOR APPLICATION NUMBER: 09/548,871
: PRIOR FILING DATE: 2000-03-30
: PRIOR APPLICATION NUMBER: 60/2126,910
: PRIOR FILING DATE: 1999-03-30
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 1164
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Described in an artificial sequence
: OTHER INFORMATION: Synthetic construct
US-10-100-745-3
```

Query Match

```
1.18: Score 17; DB 12; Length 1164,
Best Local Similarity 100.0%; Pred. No. 78;
Matches 17; Conservative 0; Mismatches 0;
```

QY 1354 CTTCGAGCTGCGAGCA 1370

DB 992 CTTCGAGCTGCGAGCA 1008

RESULT 37

```
US-09-734-032-1
: Sequence 1, Application US/09734032
: Patent No. US20020103116A1
: GENERAL INFORMATION:
: APPLICANT: Wei et al
: TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEI
: TITLE OF INVENTION: ACID MULTIPHILIC EMPLOYING HUMAN KINASE
: FILE REFERENCE: C000536
: CURRENT APPLICATION NUMBER: 09/734,032
: PRIOR FILING DATE: 2001-08-16
: PRIOR APPLICATION NUMBER: 60/205,228
: PRIOR FILING DATE: 2000-05-17
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 1275
: TYPE: DNA
: ORGANISM: human
US-09-734-032-1
```

Query Match 1.1%; Score 17; DB 10; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GTGCAGCAGCTGGGCGCT 309
 |||
 Db 463 GTGCAGCAGCTGGGCGCT 479

RESULT 38
 US-10-016-985-1
 : Sequence 1, Application US/10016985
 : Patent No. US20020123621A1
 : GENERAL INFORMATION:
 : APPLICANT: Walke, D. Wade
 : APPLICANT: Maricar, Miranda
 : APPLICANT: Yu, Xuanhuan (Sean)
 : APPLICANT: Fiddler, Carl Johan
 : TITLE OF INVENTION: No. US-09-831-458a-25 Human Kinase and Polynucleotides
 : FILE REFERENCE: Encoding the Same
 : FILE REFERENCE: LEX-0273-USA
 : CURRENT APPLICATION NUMBER: US/10/016,985
 : PRIOR FILING DATE: 2001-12-07
 : PRIOR APPLICATION NUMBER: US 60/251,941
 : PRIOR FILING DATE: 2000-12-07
 : NUMBER OF SEQ ID NOS: 3
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 1
 : LENGTH: 1275
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-10-016-985-1

Query Match 1.1%; Score 17; DB 12; Length 1275;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GTGCAGCAGCTGGGCGCT 309
 |||
 Db 463 GTGCAGCAGCTGGGCGCT 479

RESULT 39
 US-09-820-893-28/c
 : Sequence 28, Application US/09820893
 : Patent No. US20020076705A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen, et al.
 : TITLE OF INVENTION: 31 Human Secreted Proteins
 : FILE REFERENCE: P2033P1
 : CURRENT APPLICATION NUMBER: US/09/820,893
 : PRIOR FILING DATE: 2001-03-30
 : PRIOR APPLICATION NUMBER: 59/531,119
 : PRIOR FILING DATE: 2000-03-20
 : PRIOR APPLICATION NUMBER: 60/102,895
 : PRIOR FILING DATE: 1998-10-02
 : NUMBER OF SEQ ID NOS: 140
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 28
 : LENGTH: 1378
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-820-893-28

Query Match 1.1%; Score 17; DB 10; Length 1378;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 964 CAAACTGCTGAGGAGC 980
 |||
 Db 1317 CAAACTGCTGAGGAGC 1301

RESULT 40
 US-10-016-985-4
 : Sequence 3, Application US/10016985
 : Patent No. US20020123621A1
 : GENERAL INFORMATION:
 : APPLICANT: Walke, D. Wade
 : APPLICANT: Maricar, Miranda
 : APPLICANT: Yu, Xuanhuan (Sean)
 : APPLICANT: Fiddler, Carl Johan
 : TITLE OF INVENTION: No. US-09-831-458a-25 Human Kinase and Polynucleotides
 : FILE REFERENCE: Encoding the Same
 : FILE REFERENCE: LEX-0273-USA
 : CURRENT APPLICATION NUMBER: US/10/016,985
 : PRIOR FILING DATE: 2001-12-07
 : PRIOR APPLICATION NUMBER: US 60/251,941
 : PRIOR FILING DATE: 2000-12-07
 : NUMBER OF SEQ ID NOS: 3
 : SOFTWARE: FASTSEQ for Windows Version 4.0
 : SEQ ID NO 3
 : LENGTH: 1473
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-10-016-985-4

Query Match 1.1%; Score 17; DB 12; Length 1473;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GTGCAGCAGCTGGGCGCT 309
 |||
 Db 579 GTGCAGCAGCTGGGCGCT 595

RESULT 41
 US-09-808-602-26
 : Sequence 26, Application US/09808602
 : Patent No. US20020155115A1
 : GENERAL INFORMATION:
 : APPLICANT: Vornet, Corine A
 : APPLICANT: Fernandes, Elma
 : APPLICANT: Shimkets, Richard A
 : APPLICANT: Herrman, John L
 : APPLICANT: Majumder, Kamud
 : APPLICANT: Mishra, Vishnu
 : APPLICANT: Mezes, Peter S
 : APPLICANT: Macdonald, John
 : TITLE OF INVENTION: No. US-09-808-602-26 Proteins and Nucleic Acids Encoding Same
 : FILE REFERENCE: 15966-657 CIP
 : CURRENT APPLICATION NUMBER: US/09/808,602
 : PRIOR FILING DATE: 2001-03-14
 : PRIOR APPLICATION NUMBER: 69/800,198
 : PRIOR FILING DATE: 2001-03-05
 : PRIOR APPLICATION NUMBER: 50/180,590
 : PRIOR FILING DATE: 2000-03-04
 : NUMBER OF SEQ ID NOS: 114
 : SOFTWARE: Patent In Ver. 2.1
 : SEQ ID NO 26
 : LENGTH: 1498
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-808-602-26

Query Match 1.1%; Score 17; DB 9; Length 1498;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 243 TGACATCTCCGCGAG 259
 |||
 Db 1034 TGACATCTCCGCGAG 1050

RESULT 42
 US-09-848-726-1/c


```
Sequence 1, Application US/09848726
Patent No. US20020102667A1
GENERAL INFORMATION:
APPLICANT: NAMPOOHIRI, Madhavan
FILE OF INVENTION: No. US20020102667A1 Nucleotide Sequences Coding for the cbs Gen
CURRENT FILING DATE: 032301 WD 1191
CURRENT APPLICATION NUMBER: US/09/848,726
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1850
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURES:
NAME/KEY: CDS
LOCATION: (154)..(1653)
OTHER INFORMATION:
FEATURE:
NAME/KEY: RBS
LOCATION: (138)..(146)
OTHER INFORMATION:
US-09-848-726-1
```

```
Query Match
Best Local Similarity 100.0%; Score 17; DB 10; Length 1850;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1232 TTCAGACGCAATGCT 1248
|||||
DB 1281 TTCAGACGCAATGCT 1295
```

```
RESULT 43
US-09-801-220-3
Sequence 3, Application US/09801220
Patent No. US20020173020A1
GENERAL INFORMATION:
APPLICANT: Meyers, Rachel A.
APPLICANT: Williamson, Mark
TITLE OF INVENTION: 26886, A NOVEL CARBONITRILE ACYLTRANSFERASE
FILE REFERENCE: 10448-026001
CURRENT APPLICATION NUMBER: US/09/801,220
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 60/187,456
PRIOR FILING DATE: 2000-03-07
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 2412
TYPE: DNA
ORGANISM: Homo sapiens
US-09-801-220-3
```

```
Query Match
Best Local Similarity 100.0%; Score 17; DB 9; Length 2412;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 508 GCTGAGGAGATCTACG 524
|||||
DB 81 GCTGAGGAGATCTACG 97
```

```
RESULT 44
US-09-816-094-1
Sequence 1, Application US/09816094
Patent No. US20020064851A1
GENERAL INFORMATION:
APPLICANT: Wei, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
TITLE OF INVENTION: THEREOF
```

```
FILE REFERENCE: C1000536-CIP
CURRENT APPLICATION NUMBER: US/09/816,094
CURRENT FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2598
TYPE: DNA
ORGANISM: Human
US-09-816-094-1
```

```
Query Match
Best Local Similarity 100.0%; Score 17; DB 10; Length 2598;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 293 GTGACGACATGGGCT 309
|||||
DB 1236 GTGACGACATGGGCT 1252
```

```
RESULT 45
US-09-808-602-107
Sequence 107, Application US/09808602
Patent No. US20020155115A1
GENERAL INFORMATION:
APPLICANT: Verneil, Corine A
APPLICANT: Fernandes, Elma
APPLICANT: Shinkels, Richard A
APPLICANT: Herrman, John I.
APPLICANT: Majumder, Kundu
APPLICANT: Mishra, Vishnu
APPLICANT: Mezes, Peter S
APPLICANT: Macdougall, John
TITLE OF INVENTION: NO. US20020155115A1 Proteins and Nucleic Acid Sequences
FILE REFERENCE: 15966-697 CIP
CURRENT APPLICATION NUMBER: US/09/808,602
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 09/200,198
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: 60/186,596
PRIOR FILING DATE: 2000-03-03
NUMBER OF SEQ ID NOS: 114
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 107
LENGTH: 2664
TYPE: DNA
ORGANISM: Homo sapiens
US-09-808-602-107
```

```
Query Match
Best Local Similarity 100.0%; Score 17; DB 9; Length 2664;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 243 TGAACATCTGGGCGAC 259
|||||
DB 672 TGAACATCTGGGCGAC 688
```

```
Search completed: December 7, 2002, 08:48:09
Job time : 569 secs
```


[illegible][illegible]

[illegible][illegible]

SEQUENCE, 32 unordered pieces.
 AC024700
 VERSION
 AC024700.4 01:0856979
 HITS: HITS_PHASE1: HITS_PHASE1
 HITS: HITS_PHASE1: HITS_PHASE1
 SOURCE
 Homo sapiens.
 ORGANISM
 Homo sapiens.
 Mammalia: Eutheria, Chordata, Craniata, Vertebrata, Eumetazoa: Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
 1 (bases 1 to 190628)
 REFERENCE
 Waterston, R.H.
 TITLE
 The sequence of Homo sapiens clone
 JOURNAL
 2 (bases 1 to 190628)
 AUTHORS
 Waterston, R.H.
 TITLE
 Direct Submission
 Submitted (01-MAR-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 COMMENT
 On Jan 16, 2000 this sequence version replaced gi:7523915.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Project information
 Center project name: H.NH0042118
 ----- Summary Statistics -----
 Sequencing vector: M13: 100%
 Sequencing vector: plasmid: 0%
 Chemistry: Dye-primer ET, 100% of reads
 Chemistry: Dye-terminator Big Dye, 0% of reads
 Assembly program: Phrap: version 0.990319
 Consensus quality: 173687 bases at least Q40
 Consensus quality: 181172 bases at least Q20
 Consensus quality: 181172 bases at least Q20
 Insert size: 175000; agarose-IP
 Insert size: 187528; sum-of-contigs
 Quality coverage: 5.24 in Q20 bases; sum-of-contigs
 Quality coverage: 5.00 in Q20 bases; sum-of-contigs
 ----- NOTE: This is a 'working draft' sequence. It currently
 consists of 32 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available, and the accession number will
 be preserved. -----
 1 1244: contig of 1244 bp in length
 1245 1344: gap of unknown length
 1345 2396: contig of 1052 bp in length
 2397 2406: gap of unknown length
 2407 2497 3685: contig of 1089 bp in length
 3686 3685: gap of unknown length
 3686 5186: contig of 1473 bp in length
 5187 5298: gap of unknown length
 5299 6259: contig of 1359 bp in length
 6260 6724: gap of unknown length
 6725 8401: contig of 1677 bp in length
 8402 8501: gap of unknown length
 8502 10172: contig of 1671 bp in length
 10173 10272: gap of unknown length
 10273 12510: contig of 2238 bp in length
 12511 12610: gap of unknown length
 12611 14735: contig of 2125 bp in length
 14736 14835: gap of unknown length
 14836 14444: contig of 1608 bp in length
 14445 16543: gap of unknown length
 16544 19269: contig of 3726 bp in length
 19270 19369: gap of unknown length
 19370 21690: contig of 2321 bp in length
 21691 21790: gap of unknown length
 21791 25543: contig of 3753 bp in length

25544 25643: gap of unknown length
 25644 28328: contig of 2585 bp in length
 28329 28328: gap of unknown length
 28329 31649: contig of 3311 bp in length
 31650 31739: gap of unknown length
 31740 34456: contig of 4727 bp in length
 34457 36766: gap of unknown length
 36767 39593: contig of 1027 bp in length
 39594 39693: gap of unknown length
 39694 46098: contig of 6405 bp in length
 46099 46198: gap of unknown length
 46199 50494: contig of 4296 bp in length
 50495 50595: gap of unknown length
 50596 56467: contig of 5773 bp in length
 56468 56467: gap of unknown length
 56468 61903: contig of 5430 bp in length
 61904 62003: gap of unknown length
 62004 67159: contig of 5156 bp in length
 67160 67260 74171: contig of 6912 bp in length
 74172 74271: gap of unknown length
 74272 80579: contig of 6308 bp in length
 80580 80679: gap of unknown length
 80680 88540: contig of 7861 bp in length
 88541 88540: gap of unknown length
 88541 95619: contig of 6878 bp in length
 95619 95619: gap of unknown length
 95619 102484: contig of 6606 bp in length
 102485 102584: gap of unknown length
 102585 109468: contig of 6484 bp in length
 109469 109568: gap of unknown length
 109569 130222: contig of 20654 bp in length
 130223 130323: gap of unknown length
 130324 130323: gap of unknown length
 130324 154893: contig of 24567 bp in length
 154894 154989: gap of unknown length
 154990 171709: contig of 16720 bp in length
 171710 171809: gap of unknown length
 171810 171810: contig of 18819 bp in length.
 ----- FEATURES -----
 SOURCE
 1. 190628
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /colono="RP11-42018"
 1. 1244
 /note="assembly_name:Contig15"
 2445 2496
 /note="assembly_name:Contig16"
 2497 3385
 /note="assembly_name:Contig17"
 3686 5158
 /note="assembly_name:Contig20"
 5259 6624
 /note="assembly_name:Contig23"
 6725 8401
 /note="assembly_name:Contig24"
 8502 10172
 /note="assembly_name:Contig25"
 10273 12510
 /note="assembly_name:Contig29"
 12611 14735
 /note="assembly_name:Contig30"
 14836 16443
 /note="assembly_name:Contig31"
 16544 19269
 /note="assembly_name:Contig32"
 19370 21690
 /note="assembly_name:Contig33"
 21791 25543
 /note="assembly_name:Contig34"
 25644 28328
 /note="assembly_name:Contig35"
 28329 31649

BASE COUNT 310 a 315 c 346 g 226 t
ORIGIN

Query Match 7.2%, Score 111; DB 9; Length 1197;
Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACCAAGACCTGACCTTAAAGCTGCTAGTGGGAGAGCTCTGAGAGAAATCCAG 508
|||||
DB 229 ATCTACCAAGACCTGACCTTAAAGCTGCTAGTGGGAGAGCTCTGAGAGAAATCCAG 288
|||||

QY 509 CTGACAGACATCTACAGACATCTACCTGACCTGACCTGACCTGACCTGACCTGAC 339
|||||

RESULT 25
AK166105 1212 bp DNA linear PAT 17-OCT-2001
LOCUS
DEFINITION Sequence 9 from patent US 6280953.
ACCESSION AP166105
VERSION AP166105.1 GI:16241277
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Messier W. and Sikela J. M.
TITLE Methods to identify polynucleotide and polypeptide sequences which may be associated with physiological and medical conditions
JOURNAL Patent: US 6280953 A 9 28-AUG-2001
FEATURES
Source 1..1212
Location/Qualifiers
/organism="unknown"

BASE COUNT 315 a 316 c 350 g 230 t 1 others
ORIGIN

Query Match 7.2%; Score 111; DB 6; Length 1212;
Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACCAAGACCTGACCTTAAAGCTGCTAGTGGGAGAGCTCTGAGAGAAATCCAG 508
|||||
DB 229 ATCTACCAAGACCTGACCTTAAAGCTGCTAGTGGGAGAGCTCTGAGAGAAATCCAG 288
|||||

QY 509 CTGACAGACATCTACAGACATCTACCTGACCTGACCTGACCTGACCTGACCTGAC 339
|||||

RESULT 26
AK166106 1212 bp DNA linear PAT 17-OCT-2001
LOCUS
DEFINITION Sequence 10 from patent US 6280953.
ACCESSION AP166106
VERSION AP166106.1 GI:16241278
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Messier W. and Sikela J. M.
TITLE Methods to identify polynucleotide and polypeptide sequences which may be associated with physiological and medical conditions
JOURNAL Patent: US 6280953 A 10 28-AUG-2001
FEATURES
Source 1..1212
Location/Qualifiers
/organism="unknown"

BASE COUNT 315 a 316 c 350 g 230 t 1 others
ORIGIN

Query Match 7.2%; Score 111; DB 6; Length 1212;

Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACCAAGACCTGACCTTAAAGCTGCTAGTGGGAGAGCTCTGAGAGAAATCCAG 508
|||||
DB 229 ATCTACCAAGACCTGACCTTAAAGCTGCTAGTGGGAGAGCTCTGAGAGAAATCCAG 288
|||||

QY 509 CTGACAGACATCTACAGACATCTACCTGACCTGACCTGACCTGACCTGACCTGAC 339
|||||

RESULT 27
AX342003 1212 bp DNA linear PAT 10-JAN-2002
LOCUS
DEFINITION Sequence 9 from Patent WO0196603.
ACCESSION AX342003
VERSION AX342003.1 GI:18137981
KEYWORDS
SOURCE Human.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Messier W. and Sikela J. M.
TITLE Methods to identify polynucleotide and polypeptide sequences which may be associated with physiological and medical conditions
JOURNAL Patent: WO 0196603 A 9 20-JAN-2001
FEATURES
Source 1..1212
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 314 a 321 c 348 g 229 t
ORIGIN

Query Match 7.2%; Score 111; DB 6; Length 1212;
Best Local Similarity 100.0%; Pred. No. 4.4e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACCAAGACCTGACCTTAAAGCTGCTAGTGGGAGAGCTCTGAGAGAAATCCAG 508
|||||
DB 229 ATCTACCAAGACCTGACCTTAAAGCTGCTAGTGGGAGAGCTCTGAGAGAAATCCAG 288
|||||

QY 509 CTGACAGACATCTACAGACATCTACCTGACCTGACCTGACCTGACCTGACCTGAC 339
|||||

RESULT 28
AX342004 1212 bp DNA linear PAT 10-JAN-2002
LOCUS
DEFINITION Sequence 10 from Patent WO0196603.
ACCESSION AX342004
VERSION AX342004.1 GI:18137982
KEYWORDS
SOURCE Chimpanzee.
ORGANISM Pan troglodytes.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Messier W. and Sikela J. M.
TITLE Methods to identify polynucleotide and polypeptide sequences which may be associated with physiological and medical conditions
JOURNAL Patent: WO 0196603 A 10 20-DEC-2001
FEATURES
Source 1..1212
Location/Qualifiers
/organism="Pan troglodytes"
/db_xref="taxon:9598"

BASE COUNT 315 a 316 c 350 g 230 t 1 others
ORIGIN

Query Match 7.28; Score 111; DB 6; Length 1212;
 Best Local Similarity 100.0%; Pred. No. 4.4e-52;
 Matches 111; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

449 ATCTACGAGACCTGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAGAAATGCAAG 508
 DB 229 ATCTACGAGACCTGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAGAAATGCAAG 288

509 CTGTAGAGATCTTACGAGAGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAG 559
 DB 289 CTGTAGAGATCTTACGAGAGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAG 339

RESULT 29
 AF391085 1212 bp mRNA linear FRI 13-NOV-2001
 LOCUS Pan troglodytes dendritic cell-specific ICAM-3 grabbing
 DEFINITION non-integrin (DC-SIGN) mRNA, partial cds.
 ACCESSION AF391085
 VERSION AF391085.1 GI:15420781
 KEYWORDS
 SOURCE Pan troglodytes
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan;
 1 (bases 1 to 1212)
 deijenbeek, T.B., Koopman, G., van Duinhoven, G.C.F., van Vliet, S.J.,
 van Schijndel, A.C., Engering, A., Henning, T.I. and van Kooyk, Y.
 Rhesus macaque and chimpanzee DC-SIGN act as HIV/SIV gp120
 trans-receptors, similar to human DC-SIGN
 Immunol. Lett. 79 (1-2), 101-107 (2001)
 2 (bases 1 to 1212)
 deijenbeek, T.B.H., van Duinhoven, G.C.F. and Koopman, G.
 Direct Submission
 Submitted (07-JUN-2001) Molecular Cell Biology, Free University
 Medical Center Amsterdam, v.d. Boelehorststraat 9, Amsterdam 1081
 BT, The Netherlands

FEATURES
 Source
 Location/Qualifiers
 1..1212
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /cell_type="monocyte-derived dendritic cells"
 1..1212
 /gene="DC-SIGN"
 1..1212
 /note="DC-SIGN"
 /note="HIV-1 trans-receptor: CD209"
 /product="dendritic cell-specific ICAM-3 grabbing
 non-integrin"
 /protein_id="AAK97458.1"
 /db_xref="GI:15420782"
 /translation="MSDSKRPRLQQLLEPQIPGLPFPNPGVKSAGTIGHPYV
 IQLSTLIAGLVQSVKPSISOESKQVLYQNLQILAAVGLSEKSLQEIYQ
 ELTQLKAAGELPEKSKQEIYQELTRKAAGELPEKSKQEIYQELTRKAAGEL
 PEKSKQEIYQELTRKAAGELPEKSKQEIYQELTRKAAGELPEKSKQEIYQ
 LTRKAAGELPEKSKQEIYQELTRKAAGELPEKSKQEIYQELTRKAAGELPEKSKQEIYQ
 WHDSTACKEVQAQIVIKSAFQNFLLQSSSRPFWGLSDINFGCTQWWDGSP
 LUPSTNQTNRGEPNNVGRFPAPFSPGMMWTKKNI AKFWLTKKSAASDRDEQFL
 SPAPFPNPPPP"

BASE COUNT 315 a 316 c 351 g 230 t
 ORIGIN

Query Match 7.28; Score 111; DB 9; Length 1212;
 Best Local Similarity 100.0%; Pred. No. 4.4e-52;
 Matches 111; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

449 ATCTACGAGACCTGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAGAAATGCAAG 508
 DB 229 ATCTACGAGACCTGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAGAAATGCAAG 288

509 CTGTAGAGATCTTACGAGAGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAG 559
 DB 289 CTGTAGAGATCTTACGAGAGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAG 339

RESULT 30
 AX039403 1215 bp DNA linear
 LOCUS Sequence 1 from Patient W-9063251.
 DEFINITION AX039403
 ACCESSION AX039403
 VERSION AX039403.1 GI:11229473
 KEYWORDS
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;
 1 (bases 1 to 1215)
 Pijbor, G., Geijtenbeek, T.B., van Kooyk, Y., and
 Composition and method for modulating dendritic
 cell
 Patent: WO 0063251-A 1 26-NOV-2000;
 Katholieke Universiteit Nijmegen (NL)

FEATURES
 Source
 Location/Qualifiers
 1..1215
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Human; DC-SIGN"
 BASE COUNT 315 a 321 c 349 g 230 t
 ORIGIN

Query Match 7.28; Score 111; DB 6; Length 1212;
 Best Local Similarity 100.0%; Pred. No. 4.4e-52;
 Matches 111; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

449 ATCTACGAGACCTGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAGAAATGCAAG 508
 DB 229 ATCTACGAGACCTGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAGAAATGCAAG 288

509 CTGTAGAGATCTTACGAGAGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAG 559
 DB 289 CTGTAGAGATCTTACGAGAGACCTGATTAAGGCTGAGTGGTGTGCTCTCAGCAAG 339

RESULT 31
 AY042221 1215 bp mRNA linear
 LOCUS Homo sapiens mdc-SIGN1A type 1 isoform (CD209) mdc
 DEFINITION AY042221
 ACCESSION AY042221
 VERSION AY042221.1 GI:15281072
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;
 1 (bases 1 to 1215)
 Mummidi, S., Catano, G., Lam, L., Hoefle, A., Telles,
 Jimenez, F., Ahuja, S.S., and Ahuja, S.K.
 Extensive repertoire of membrane bound and soluble
 cell specific ICAM-3 grabbing nonintegrin 1 (DC-SIGN)
 isoforms: Inter individual variation in expression
 transcripts
 1 Biol Chem 276 (45), 43194-43212 (2001)

JOURNAL MEDLINE
 21413847
 11337487
 2 (bases 1 to 1215)
 Mummidi, S., and Ahuja, S.K.
 Direct Submission
 Submitted (18-JUN-2001) Department of Medicine,
 Veterans Health Care System and University of Texas
 Center, 7703, Floyd Curl Drive, San Antonio, TX
 78229

FEATURES
 Source
 Location/Qualifiers
 1..1215

ORIGIN

Query Match 7.2%: Score 111; DB 6; Length 1312;
 Best Local Similarity 100.0%; Pred. No. 4,4e-52;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 508
 |||||||
 DB 270 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 329
 |||||||
 QY 509 CTGACGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGAG 559
 |||||||
 DB 340 CTGACGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGAG 490
 |||||||

RESULT 14
 HOMOLECTIN
 LOCUS 1312 bp mRNA linear PR: 19-JUN-2000
 DEFINITION Homo sapiens membrane-associated lectin type-C mRNA, complete cds
 ACCESSION M98457
 VERSION M98457.1 GI:187115
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 FEATURES

1 (bases 1 to 1312)
 Sequence and expression of a membrane-associated C-type lectin that exhibits CD4-independent binding of human immunodeficiency virus envelope glycoprotein gp120
 Proc. Natl. Acad. Sci. U.S.A. 89 (17), 8356-8360 (1992)

LOCATION/QUALIFIERS
 1. 1312
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue-type="placenta"
 42..1256
 /codon_start=1
 /product="membrane-associated lectin type-C"
 /protein_id="AAE77072.1"
 /db_xref="GI:8572543"

BASE COUNT 336 a 348 c 372 g 256 t
 ORIGIN
 Query Match 7.2%: Score 111; DB 6; Length 1312;
 Best Local Similarity 100.0%; Pred. No. 4,4e-52;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 508
 |||||||
 DB 270 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 329
 |||||||
 QY 509 CTGACGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGAG 559
 |||||||
 DB 340 CTGACGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGAG 490
 |||||||

RESULT 15
 AY042229 1315 bp mRNA linear PR: 28-AUG-2001
 LOCUS
 DEFINITION Homo sapiens MDC-SIGNIB type I isoform (CD209) mRNA, complete cds,
 alternatively spliced.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

FEATURES
 SOURCE

1. 1315
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19p11.4"
 1..1315
 /gene="CD209"
 1..1126
 /note="CD209"
 /note="putative open reading frame; encoding protein MDC-SIGNIB"
 /evidence=not_experimental
 101..1315
 /gene="CD209"
 /note="Similar to the nucleotide sequence Genbank Accession Number AF008812; also /product="MDC-SIGNIB type I isoform"
 /protein_id="AAK91854.1"
 /db_xref="GI:15281089"

BASE COUNT 332 a 345 c 496 g 252 t
 ORIGIN
 Query Match 7.2%: Score 111; DB 6; Length
 Best Local Similarity 100.0%; Pred. No. 4,4e-52;
 Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 508
 |||||||
 DB 329 ATCTACGAGAGCTGACCTGATTAAGCTGATGCTGATGCTGATGAGGAATCCAG 329
 |||||||
 QY 509 CTGACGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGAG 559
 |||||||
 DB 389 CTGACGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGAG 490
 |||||||

RESULT 16
 AF290886 4266 bp mRNA linear PR: 28-AUG-2001
 LOCUS
 DEFINITION Homo sapiens DC-SIGN mRNA, complete cds.

Db 1428 CTGACAGAGATCTACGAGAGCTGACCACTTAATAGTCTCAGTGGCTGAG 1478

RESULT 38
AC008812/c 143619 bp DNA linear PK1 28-JUL-2000

LOCUS AC008812
DEFINITION Homo sapiens chromosome 19 clone CTD-2102F19, complete sequence.
AC008812
VERSION AC008812.7 GI:9558577
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT 2 (bases 1 to 143619)
Direct Submission
DOI Joint Genome Institute.
JOURNAL Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 143619)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Submitted (28-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 28, 2000 this sequence version replaced a previous version.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 3.8.
Location/Qualifiers
1. 143619
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-2102F19"

BASE COUNT 36628 a 45485 c 36414 g 36192 t

ORIGIN

Query Match 7.2%; Score 111; DB 9; Length 143619.
Best Local Similarity 100.0%; Pred. No. 3 9e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 449 ATCTACGAGAGCTGACGAGCTTAAAGCTGACGTGGTGGTCTCTCAGAGAAATCCAAAG 508
|||||
Db 49148 ATCTACGAGAGCTGACGAGCTTAAAGCTGACGTGGTGGTCTCTCAGAGAAATCCAAAG 49089
|||||

QY 509 CTGACAGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAG 559
|||||
Db 49088 CTGACAGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAG 49038
|||||

RESULT 39
AC008763/c 177962 bp DNA linear PK1 16-JUL-2000

LOCUS AC008763
DEFINITION Homo sapiens chromosome 19 clone CTD-3214H19, complete sequence.
AC008763
VERSION AC008763.9 GI:21844554
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT 2 (bases 1 to 177962)

AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 177962)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Submitted (01-JUL-2002) DOE Joint Genome Institute
JOURNAL Submitted (16-JUL-2002) DOE Joint Genome Institute
4 (bases 1 to 177962)
DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Submitted (16-JUL-2002) DOE Joint Genome Institute
JOURNAL Submitted (16-JUL-2002) DOE Joint Genome Institute
On Jul 16, 2002 this sequence version replaced a previous version.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.7% of Sequence;
Estimated Total Number of Errors is 0.5.
NOTE: Break in plasmid subclone coverage 122800-123177. Unresolved repeat copies 122800-123177.
Location/Qualifiers
1. 177962
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-3214H19"
/note="NOTE: Break in plasmid subclone 122800-123177. Unresolved tandem repeat 122800-123177. Unresolved repeat copies 122800-123177."

BASE COUNT 40697 a 49263 c 46202 g 41800 t

ORIGIN

Query Match 7.2%; Score 111; DB 9; Length 177962.
Best Local Similarity 100.0%; Pred. No. 3 9e-52;
Matches 111; Conservative 0; Mismatches 0; Indels 0;

QY 449 ATCTACGAGAGCTGACGAGCTTAAAGCTGACGTGGTGGTCTCTCAGAGAAATCCAAAG 508
|||||
Db 176849 ATCTACGAGAGCTGACGAGCTTAAAGCTGACGTGGTGGTCTCTCAGAGAAATCCAAAG 508
|||||

QY 509 CTGACAGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAG 559
|||||
Db 176789 CTGACAGAGATCTACGAGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAGCTGACGAG 559
|||||

RESULT 40
AF369755 1146 bp mRNA linear PK1 16-JUL-2000

LOCUS AF369755
DEFINITION Macaca mulatta dendritic cell-specific 10A1-3 gene.
AF369755
VERSION AF369755.1 GI:16118474
KEYWORDS mRNA, complete cds.
SOURCE Macaca mulatta.
ORGANISM Macaca mulatta.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
AUTHORS Hariband, F., Pohlman, S., Sparwasser, T., Kimata, N., Haggarty, B.S., Ahmad, N., Maitalan, T., Edwards, J., Arnason, J., Reihardt, J.A., Kimata, J.T., Littman, D. and Doms, R.W.
TITLE Functional and anti-viral characterization of human plasmalogen macaque and murine 10A1-3 gene
JOURNAL J. Virol. 75 (21), 10281-10289 (2001)
MEDLINE 21465051
PUBMED 11581396


```

REFERENCE      1
AUTHORS        Messier, W. and Sikela, J.M.
TITLE          Methods to identify polynucleotide and polypeptide sequences which
               may be associated with physiological and medical conditions
JOURNAL        Patent. WO 01/66013-A 11-20-DEC-2001;
               Evolutionary Genomics, LLC (US)
FEATURES
SOURCE         1. 1212
               /organism="Gorilla gorilla"
               /db_xref="taxon:9593"
BASE COUNT     315 a 319 c 349 g 229 t
ORIGIN
Query Match   4.18; Score 63; DB 6; Length 1212;
Best Local Similarity 100.08; Pred. No. 2,7e-24;
Matches 63; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

OY 860 TCGACATCTTCGCAAGCAACCTGTTACTGATGCTAACTTCGCAAGCAACCTGCAAGCAAC
Db 778 TCGACATCTTCGCAAGCAACCTGTTACTGATGCTAACTTCGCAAGCAACCTGCAAGCAAC 817

OY 920 TGC 922
Db 818 TGC 840

RESULT 44
AF209480S1    603 bp DNA linear FRI 12 JAN 2001
LOCUS         Homo sapiens probable mannose binding C-type lectin DC-SIGNR gene,
DEFINITION    exons 1 and 2.
ACCESSION     AF209480
VERSION        AF209480.1 GI:10179612
KEYWORDS
SEGMENT       1 of 2
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiini; Homo.
REFERENCE     1 (bases 1 to 603)
AUTHORS       Sollier, E.J., Barten, R. and Trowsdale, J.
JOURNAL       DC-SIGN: a related gene, DC-SIGNR, and CD23 form a cluster on 19p13
TITLE         20432267
MEDLINE       10975799
PUBMED        2 (bases 1 to 603)
REFERENCE     Sollier, E.J.
AUTHORS       Direct Submission
JOURNAL       Submitted (26 NOV 1999) Department of Pathology, University of
               Cambridge, Tennis Court Road, Cambridge, Cambs CB2 1QP, UK
FEATURES
SOURCE        1. 603
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="19"
               /map="19p13; adjacent to CD23"
               /gene="abc"
               /number=1
               /number=2
               /number=2
               /number=2
BASE COUNT     144 a 123 c 218 g 118 t
ORIGIN
Query Match   3.88; Score 59; DB 9; Length 603;
Best Local Similarity 100.08; Pred. No. 5,7e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

OY 257 CAGCGGCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
Db 13 CAGCGGCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 71

```

```

RESULT 45
AF343727
LOCUS         1146 bp mRNA linear
DEFINITION    Macaca nemestrina dendritic cell-specific rRNA-5
ACCESSION     AF343727
VERSION        AF343727.1 GI:16118454
KEYWORDS
SOURCE        Macaca nemestrina.
ORGANISM      Macaca nemestrina
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecoidea; Macaca.
REFERENCE     1 (bases 1 to 1146)
AUTHORS       Baribaud, F., Pohlmann, S., Sparwasser, T., Yu, K.,
               Haegarty, B.S., Ahmad, N., Marzarian, T., Edwards, T.,
               Arnason, J., Reinhardt, T.A., Kimata, J.T., Littman, D.
               and Doms, R.W.
TITLE         Functional and antigenic characterization of human
               dendritic and antigen-presenting cells from
               plaited macaque, and murine DC-SIGN
JOURNAL       J. Virol. 75 (21), 10281-10289 (2001)
MEDLINE       21465051
PUBMED        11581396
REFERENCE     2 (bases 1 to 1146)
AUTHORS       Kimata, J.T. and Yu, Kimata, M.T.
TITLE         Direct Submission
JOURNAL       Submitted (29-JAN-2001) Virology & Immunology, School
               for Biomedical Research, c/o. Box 760549, San Antonio
               78245-0549, USA
FEATURES
SOURCE        Location/Qualifiers
               1. 1146
               /organism="Macaca nemestrina"
               /db_xref="taxon:9545"
               /cell_type="peripheral blood-derived dendritic
               cell"
               /note="membrane-associated mannose binding
               protein"
               /codon_start=1
               /product="dendritic cell-specific rRNA-5
               nonIntegrin"
               /protein_id="AA114428.1"
               /db_xref="GI:16118455"
               /translation="MSDSKEPRIOGQDILJEEBOLGVRP
               LQISFTLLAGLIVQSKSPSSISGOSKODATYONITQIP
               ELRLKAAYELPEKSKOQETFEELRLKAAYGELPEKSK
               PEKSKOQETFEELRLKAAYGELPEKSKOQETFEELRLKAAY
               LQILKAVERLQHPCHWEMFFQNGCTFMNSORNMHS117
               QNELQSSRSNMFIMWJSDINHECTWGVGSPILSPK
               EFGNWNIDKCMIAFWICKKSAASGSDERKLSPAP118
BASE COUNT     296 a 313 c 324 g 214 t
ORIGIN
Query Match   3.78; Score 57; DB 9; Length 1146;
Best Local Similarity 100.08; Pred. No. 8,2e-21;
Matches 57; Conservative 0; Mismatches 0; Indels 0;

OY 572 TCGAAGCTGTAGGAGATCTACGAGAGTACGAGAGTACGAGAGTACGAGAGTACGAGAGT
Db 421 TCGAAGCTGTAGGAGATCTACGAGAGTACGAGAGTACGAGAGTACGAGAGTACGAGAGT 71

```

Search completed: December 7, 2002, 10:41:02
 Job time : 3787 secs

RESULT 2
AAV84361
ID AAV84361 standard; cDNA to mRNA; 1643 BP.
XX
AC AAV84361:
XX
DT 30-MAR-1999 (first entry)
XX
DE Human liver cDNA clone HP01347.
XX
KW Transmembrane protein; HP01347; human; lectin; receptor; liver; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 25..915
FT /tag= a
FT /note= "cDNA comprising the coding region (minus
FT the stop codon) is claimed (claim 3)"
XX
PN W0985508-A2.
XX
PD 10-DEC-1998.
XX
PE 03-JUN-1998; 98MO-JP02445.
XX
PK 04-JUN-1997; 97JP-014448.
XX
PA (PRO1-) PROTEGENE INC.
PA (SAGA) SAGAMI CHEM RES. CENTRE.
XX
PI Kato S, Sekine S, Yamaguchi T;
XX
DR WPI: 1999-045730/04.
DR P-PSDB: AAM88493.
XX
PT New human proteins containing transmembrane domains and their
PT encoding sequences - useful in the preparation of antibodies and
PT large-scale protein production, gene diagnosis, and gene therapy
XX
PS Claim 4: Page 125-126; 178pp; English.
XX
XX This is the nucleotide sequence of cDNA clone HP01347, which
XX includes a coding region (also claimed) for a novel human
XX transmembrane protein (see AAM88493). The clone was isolated from a
XX liver cDNA library using a signal sequence detection method and by
XX protein synthesis by in vitro translation. The encoded protein
XX is characterised as a type-II membrane protein having a single
XX N-terminal transmembrane domain. It has 85.5% homology over 284
XX amino acid residues to the human HIV envelope glycoprotein gp120-
XX binding C-type lectin (a CD4-independent HIV receptor) and may
XX function as a receptor on the membrane surface. The invention
XX provides nucleotide sequences (see AAV84359-76) coding for 13
XX transmembrane proteins (see AAM88491-508), vectors containing such
XX polynucleotides, and eukaryotic cells containing the vectors. The
XX proteins can be used as antigens or as compositions in the
XX preparation of antibodies against the proteins. The polynucleotides
XX can be used as probes for gene diagnosis, and as gene sources for
XX gene therapy and large-scale production of proteins encoded by the
XX cDNA. The host cells are used for the detection of ligands
XX corresponding to the expressed proteins, and the screening of low
XX mol.wt. medicines.
XX
XX Sequence 1643 BP; 388 A; 429 C; 440 G; 386 T; 0 other:

Query Match 33.6%; Score 518; DB 20; Length 1643;
Best Local Similarity 100.0%; Pred. No. 6,5e-246;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 257 CAGCGGAAACATGAGTCTCAAGGAAACCAAGGAGTGGAGCTCTCTGGGG 316
DB 13 CAGCGGAAACATGAGTCTCAAGGAAACCAAGGAGTGGAGCTCTCTGGGG 72

UY 317 TGCTTGGCCATGAGCGGCTGATGTGTAACCTCTCTCTTCATGCTCTT
DB 73 TGCTTGGCCATGAGCGGCTGATGTGTAACCTCTCTCTTCATGCTCTT
UY 377 CTGTTGGCCATGCTCTCAAGGAAACCAAGGAGTGGAGCTCTCTGGGG 316
DB 133 CTGTTGGCCATGCTCTCAAGGAAACCAAGGAGTGGAGCTCTCTGGGG 72
UY 437 GAGCAGAGCGCATCTACCAAGGAAACCAAGGAGTGGAGCTCTCTGGGG 316
DB 133 GAGCAGAGCGCATCTACCAAGGAAACCAAGGAGTGGAGCTCTCTGGGG 72
UY 497 GAGAAATCTAAATCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 316
DB 253 GAGAAATCTAAATCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 72
UY 557 GAGTTGGCCAGGAAATCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 316
DB 313 GAGTTGGCCAGGAAATCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 72
UY 617 GAGTTGGCCAGGAAATCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 316
DB 373 GAGTTGGCCAGGAAATCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 72
UY 677 CTGAGGAGTCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 316
DB 433 CTGAGGAGTCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 72
UY 737 CTGAGGAGTCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 316
DB 493 CTGAGGAGTCTGAGGAAACCAAGGAGTGGAGCTCTCTGGGG 72
RESULT 3
AA165469
ID AA165469 standard; DNA; 1643 BP.
XX
AC AA165469:
XX
DT 10-DEC-2001 (first entry)
XX
DE DNA encoding a splice variant of human DC-SIGN.
XX
KW Human; receptor; DC-SIGN; dendritic cell; T lymphocyte;
KW gp120; C-type lectin; ICAM3; HIV entry; T cell; macrophage;
KW HIV infection; splice variant; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 25..915
FT /tag= a
FT /product= "DC-SIGN splice variant"
XX
PN W0200164752-A2.
XX
PD 07-SEP-2001.
XX
PE 28-FEB-2001; 2001WO-050672.
XX
PK 02-MAR-2000; 2000US-0517605.
XX
PA (UYNV) UNIV NEW YORK STATE.
PA (UYNV) UNIV NIMMGEN.
XX
PI Littman DR, Kwon D, Van Koeck Y, Geijtenbeek T;
XX
DR WPI: 2001-602565/68.
DR P-PSDB: AAG79093.
XX
PT An antibody for the treatment or prevention of HIV-infection
PT is provided which binds to DC-SIGN or is exposed up
PT on DC-SIGN due to concomitant conformational change.

RESULT 8
ABN94615/c

ID ABN94615 standard; DNA; 506 BP.

AC ABN94615;

DT 13-AUG-2002 (first entry)

DE Gene #1113 used to diagnose liver cancer.

XX Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX metastatic liver tumor; cytostatic; expression profile; disease state;
XX disease progression; drug toxicity; drug efficacy; drug metabolism.

OS Homo sapiens.

PN W020029103 A2.

PD 11-APR-2002.

PE 02-OCT-2001; 2001WO-0830589.

PR 02-OCT-2000; 2000US-237054P.

PA (GENE-) GENE LOCUS INC

PI Horne D, Alvares C, Peres Da-Silva S, Winkley TG.

PR 2002-426119/45.

PT Diagnosing and detecting the progression of liver cancer.

PT hepatocellular carcinoma or metastatic liver tumor in a patient,

PT involves detecting the level of expression of two or more genes in a

XX liver tissue sample.

XX Claim 1: SEQ ID NO 1113; 298pp; English.

XX The invention relates to a novel method for diagnosing and detecting the
XX progression of liver cancer, hepatocellular carcinoma or metastatic liver
XX tumor in a patient, and differentiating metastatic liver cancer from
XX hepatocellular carcinoma in a patient, involving detecting the level of
XX expression of two or more genes represented in ABN93503-ABN97455 in a
XX tissue sample. The method of the invention has hepatotropic, and
XX cytostatic activity. The method is useful for diagnosing and detecting
XX the progression of liver cancer, hepatocellular carcinoma and metastatic
XX liver carcinoma in a patient. The method is useful for identifying
XX expression profiles which serve as useful diagnostic markers as well as
XX markers that can be used to monitor disease states, disease progression,
XX drug toxicity, drug efficacy and drug metabolism.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at http://wipo.int/pub/published_pat_sequences.

XX Sequence 506 BP; 165 A; 112 C; 127 G; 102 T; 0 other;

Query Match 14.3%; Score 220; DB 24; Length 506;

Best Local Similarity 100.0%; Pred No 1 9e+98;

Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1318 CTTTGACTGGGGAGCTTCTGTTGAAGAGGACTCATCTTGAGGCTGGAGGACGACGGA 1377

DB 220 CTTTGACTGGGGAGCTTCTGTTGAAGAGGACTCATCTTGAGGCTGGAGGACGACGGA 161

UY 1378 ATTAATTCGCCAGTCAGCAATGCAATGATAGAGAGAGATGAGAGGAGGAGGAGGAGG 1437

DB 160 ATTAATTCGCCAGTCAGCAATGCAATGAGAGAGAGAGGAGGAGGAGGAGGAGGAGG 101

UY 1418 TTCACGCCGACATCTTCTTGTCTATACATGCTTTCATTTGGCTTTTCTGAGTTS 1497

DB 100 TTCACGCCGACATCTTCTTGTCTATACATGCTTTCATTTGGCTTTTCTGAGTTS 41

UY 1498 TAGCCTTATATAAAGTGTAATGTTGTAATGCAAAA 1537

DB 40 TAGCCTTATATAAAGTGTAATGTTGTAATGCAAAA 1

RESULT 9

ID ABA60904 standard; DNA; 592 BP.

AC ABA60904;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #9269

XX Human; foetal liver; gene expression; single exon nucleic

XX acid probe.

OS Homo sapiens.

PN W0200157277 A2.

PD 09-AUG-2001.

PE 30-JAN-2001; 2001WO-0509669.

PR 04-FEB-2000; 2000US-0196332.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0634506.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0235359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI: 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes us

XX analyzing gene expression in human foetal liver.

XX Claim 1: SEQ ID NO 9269; 639pp - sequence listing; Eng;

XX The invention relates to a single exon nucleic acid probe

XX measuring human gene expression in a sample derived from

XX liver. The single exon nucleic acid probes may be used to

XX measure and displaying gene expression in samples derived

XX from foetal liver. The present sequence is a single exon nucleic

XX acid probe of the invention.

XX Note: The sequence data for this patent did not form part

XX of the printed specification, but was obtained in electronic

XX format from WIPO at http://wipo.int/pub/published_pat_sequences.

XX Sequence 592 BP; 154 A; 145 C; 168 G; 121 T; 0 other;

Query Match

Best Local Similarity 9.9%; Score 153; DB 22; Length

Matches 153; Conservative 0; Mismatches 0; Indels 0;

UY 830 GAAGGCTGTGTGCGGCACTGTGTGAAGAGGAGATGACATTTCTTGAAGAGAG 1

DB 149 GAAGGCTGTGTGCGGCACTGTGTGAAGAGAGATGACATTTCTTGAAGAGAG 200

UY 890 ATGCTTAATCTCCAGGAGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 349

DB 209 ATGCTTAATCTCCAGGAGGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 249

UY 950 CAGCTGCTGTAATCAAAACTGTGAGAGAGAG 982

DB 269 CAGCTGCTGTAATCAAAACTGTGAGAGAGAG 301

RESULT 10

```

ABA28884
ID ABA28884 standard; DNA: 592 BP.
XX
AC ABA28884;
XX
XX 23-JAN-2002 (first entry)
XX
LE Probe #7550 for gene expression analysis in human heart cell sample.
XX
KW Human: gene expression, heart, microarray, vascular system, probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
CS Homo sapiens.
XX
XX MO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001MO-US00666
XX
XX 04-FEB-2000; 2000MS-0180312
XX 26-MAY-2000; 2000MS-0297456.
XX 30-JUN-2000; 2000MS-0608408.
XX 03-AUG-2000; 2000MS-0632466.
XX 21-SEP-2000; 2000MS-0234687.
XX 27-SEP-2000; 2000MS-0236379.
XX 04-OCT-2000; 2000CB-0024262.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Park DP;
XX WPI; 2001-48889q/53
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts -
XX
XX Claim 1: SEQ ID NO 7350; 530bp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX measuring human gene expression in a sample derived from human heart. The
XX present sequence is one such probe. The probes may be used for
XX predicting, measuring and displaying gene expression in samples derived
XX from the human heart via microarrays. By measuring gene expression, the
XX probes are useful for predicting, diagnosing, grading, staging,
XX monitoring and prognosing diseases of the human heart and vascular system
XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX congenital heart disease.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WPI
XX at http://wipo.int/pub/published/seq\_sequences
XX
XX Sequence 592 BP, 154 A, 149 C, 168 G, 121 T, 0 other:
XX
Query Match 9.9%; Score 153, DB 22, Length 592;
Best Local Similarity 100.0%; Pred. No. 2.7e-65;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

AAK09194
ID AAK09194 standard; DNA: 592 BP.
XX
AC AAK09194;
XX
XX 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 9185.
XX
XX Human: brain expressed exon, gene expression analysis, probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX
XX Homo sapiens.
XX
XX MO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001MO-US00667.
XX
XX 04-FEB-2000; 2000MS-0180312.
XX 26-MAY-2000; 2000MS-0297456.
XX 30-JUN-2000; 2000MS-0608408.
XX 03-AUG-2000; 2000MS-0632466.
XX 21-SEP-2000; 2000MS-0234687.
XX 27-SEP-2000; 2000MS-0236379.
XX 04-OCT-2000; 2000CB-0024262.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Park DP;
XX WPI; 2001-48889q/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4: SEQ ID NO: 9185; 650bp + sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 592 BP, 154 A, 149 C, 168 G, 121 T, 0 other:
XX
Query Match 9.9%; Score 153, DB 22, Length 592;
Best Local Similarity 100.0%; Pred. No. 2.7e-65;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

UT 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 9640
DE
XX
XX Human: bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX Homo sapiens.
OS
XX W0200157276-A2.
XX
XX 30-JAN-2001; 2001WO-US006648
XX
XX 30-JAN-2001; 2001WO-US006648
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOL- ) MOLECULAR DYNAMICS INC.
XX
XX Penn SQ, Hanzel DK, Chen W, Rank DR;
XX
XX WPI: 2001-488900/53
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO: 9640; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in bone marrow
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention.
XX
XX Sequence 592 BP; 154 A; 149 G; 168 C; 121 T; 0 other.
XX
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 153; DB 22; Length 592;
XX Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
XX
XX QY 830 GAAGGCTGTGCGGCACTGTGCTCAAGGACTGACATTTCTTCAAGTAACCTTACTTC 889
XX |||||||
XX DB 149 GAAGGCTGTGCGGCACTGTGCTCAAGGACTGACATTTCTTCAAGTAACCTTACTTC 208
XX |||||||
XX QY 890 ATGCTACTGCTCCAGGCACTGTGCTCAAGGACTGACATTTCTTCAAGTAACCTTACTTC 949
XX |||||||
XX DB 209 ATGCTACTGCTCCAGGCACTGTGCTCAAGGACTGACATTTCTTCAAGTAACCTTACTTC 208
XX |||||||
XX QY 950 CAGCTGCTGTATATCAAAACTGTAGAGAGAG 982
XX |||||||
XX DB 269 CAGCTGCTGTATATCAAAACTGTAGAGAGAG 301
XX |||||||
XX
XX RESULT 13
XX ID AAI40799 standard; DNA: 592 BP.
XX
XX AAI40799:
XX
XX 17-OCT-2001 (first entry)
XX
XX Probe #9485 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX
XX

```

```

OS Homo sapiens.
XX
XX W0200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US006648.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOL- ) MOLECULAR DYNAMICS INC.
XX
XX Penn SQ, Hanzel DK, Chen W, Rank DR;
XX
XX WPI: 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes use
XX analyzing gene expression in human placenta.
XX
XX Claim 25; SEQ ID NO 9485; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid
XX probes which are derived from genomic sequences expressed in human
XX placenta. They can be used to measure gene expression in human
XX placenta samples, which may enable the improved diagnosis and
XX treatment of disorders.
XX
XX Sequence 592 BP; 154 A; 149 G; 168 C; 121 T; 0 other.
XX
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 153; DB 22; Length 592;
XX Matches 153; Conservative 0; Mismatches 0; Indels 0.
XX
XX QY 830 GAAGGCTGTGCGGCACTGTGCTCAAGGACTGACATTTCTTCAAGTAACCTTACTTC 889
XX |||||||
XX DB 149 GAAGGCTGTGCGGCACTGTGCTCAAGGACTGACATTTCTTCAAGTAACCTTACTTC 208
XX |||||||
XX QY 890 ATGCTACTGCTCCAGGCACTGTGCTCAAGGACTGACATTTCTTCAAGTAACCTTACTTC 949
XX |||||||
XX DB 209 ATGCTACTGCTCCAGGCACTGTGCTCAAGGACTGACATTTCTTCAAGTAACCTTACTTC 208
XX |||||||
XX QY 950 CAGCTGCTGTATATCAAAACTGTAGAGAGAG 982
XX |||||||
XX DB 269 CAGCTGCTGTATATCAAAACTGTAGAGAGAG 301
XX |||||||
XX
XX RESULT 14
XX ID ABS09548 standard; DNA: 592 BP.
XX
XX ABS09548:
XX
XX 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon probe from lung SEQ ID NO:
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD;
XX chronic obstructive pulmonary disease; interstitial lung
XX disease; familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Rudrak syndrome; sarcoidosis; pulmonary haemorrhage;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Kaposi's
XX sarcoma; primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
XX
XX
XX
XX
XX Homo sapiens.
XX
XX
XX

```


KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberosus sclerosis; Gaucher's disease; Niemann-pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.
 XX
 XX Homo sapiens.
 OS
 PN W02001R600R-A2
 XX
 XX 15-NOV-2001.
 PD
 XX
 PF 30-JAN-2001; 2001WO-US000665.
 XX
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-063236P.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236353P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLF-.) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PT WP1: 2002-114183/15.
 TP
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 XX
 PS Claim 4; SEQ ID NO 21954; 63pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one
 CC of 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes: the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung, measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene, a method comprising one
 CC of 12011 sequences, mentioned in the specification, or employed by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease,
 CC Niemann-pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic


```

XX PS Example 19, Fig 12, 141pp, English.
XX CC
CC CC The present sequence is that of the coding region of the chimpanzee
CC CC DC-SIGN gene. DC-SIGN is expressed on dendritic cells and is known
CC CC to provide a mechanism for transport of HIV-1 virus to the lymph
CC CC nodes. HIV-1 binds to the extracellular portion of DC-SIGN and
CC CC infects the undifferentiated T cells in the lymph nodes via their
CC CC CD4 proteins. This ultimately leads to compromise of the immune
CC CC system and to full-blown AIDS. The invention comprises a
CC CC comparative genomics approach to identify specific gene changes
CC CC responsible for differences in functions and diseases distinguishing
CC CC humans from non-humans, particularly primates including gorilla,
CC CC orangutan and especially chimpanzee. The evolutionary significance
CC CC of a nucleotide change is determined by the ratio of the
CC CC non-synonymous substitution rate (Ka) to the synonymous rate (Ks)
CC CC of the nucleotide sequence. Polynucleotide and polypeptide
CC CC sequences corresponding to evolved traits may be relevant to human
CC CC diseases or conditions such as unique or enhanced human brain
CC CC functions, longer human life spans, susceptibility or resistance to
CC CC disease, including AIDS and cancer, and aesthetic traits such as
CC CC hair growth. Ka/Ks ratios for chimpanzee/human, human/gorilla and
CC CC chimpanzee/gorilla DC-SIGN cDNA sequences are 1.3, 0.87 and 1.3,
CC CC respectively. It is theorized that chimpanzee resistance to
CC CC progression to full-blown AIDS may be due in part to inability of
CC CC HIV-1 to bind to chimpanzee DC-SIGN for transport to the lymph
CC CC nodes. After determining the 3-dimensional structure of DC-SIGN,
CC CC a rational drug design approach can be used to mimic the effects
CC CC of chimpanzee DC-SIGN without interfering with the normal functions
CC CC of human DC-SIGN.
XX SS
SS Query Match
SS Best Local Similarity 100.0%; Score 111; DB 24; Length 1212;
SS Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 449 ATCTACCGAAGCTGACCCAGCTTAAGCTGAGTGGCTCTGAGAAATCCAG 508
DB 229 ATCTACCGAAGCTGACCCAGCTTAAGCTGAGTGGCTCTGAGAAATCCAG 268
OY 509 CTGCAGGAGATCTACGAGTACCTGAGTGAAGCTGCGAGTGGGTGAG 559
DB 289 CTGCAGGAGATCTACGAGTACCTGAGTGAAGCTGCGAGTGGGTGAG 339

```

RESULT 24
AAA88740
ID AAA88740 standard; cDNA; 1215 BP.
XX AAA88740;
AC 19-FEB-2001 (first entry)
DT
XX
DE Dendritic cell specific C-type lectin DC-SIGN cDNA.
KM Dendritic cell, lectin; DC-SIGN; intercellular adhesion molecule;
KM ICAM-3 receptor; human; HIV-1; infection; autoimmune disease
KM allergy; immunotherapy; immunosuppressive; therapy; ss.
OS
XX Homo sapiens.
PN EPI046651-A1.
XX
PO 25-OCT-2000.
XX
PF 19-APR-1999; 99EP-0201234.
XX
PR 19-APR-1999; 99EP-0201234.
XX
PA (UTN1-) UNIV KONINK NIDMEGEN.
XX
PI Fiedor CG, Geltenbeek TBH, Van Kooyk Y, Torrensma P;

```

XX PS WP1: 2000-040929/63.
XX CC
XX CC P-PSDB: AAH19714.
XX CC
XX CC Modulating the immune response in an animal, useful for
XX CC immunosuppression, using a compound that binds to C-type
XX CC dendritic cells
XX PS
PS Disclosure: Page 17-18; 44pp; English.
XX CC
XX CC The present sequence is that of cDNA encoding DC-SIGN
XX CC a 44 kDa C-type lectin of human dendritic cells (DC).
XX CC novel ICAM-3 receptor that is specifically expressed by
XX CC It is involved in the initial transient DC-T-cell inter-
XX CC necessary for initiating an immune response. DC-SIGN
XX CC to bind to the HIV envelope protein gp120 and to facilitate
XX CC entry into DC. The cDNA was obtained by RT-PCR amplification
XX CC AAA88741-42) of DC RNA. The invention relates to the use
XX CC compound that binds to a C-type lectin (especially DC-SIGN)
XX CC surface of a DC, and in the preparation of a composition
XX CC modulating, especially reducing, the immune response in
XX CC in particular a human or other mammal. The composition
XX CC the interactions between a DC and a T-cell, more specifically
XX CC between a C-type lectin on the surface of a DC and an
XX CC on the surface of a T-cell. It is used to prevent/inhibit
XX CC responses to specific antigens, for inducing tolerance,
XX CC immunotherapy, for immunosuppression, for the treatment
XX CC autoimmune diseases, the treatment of allergy, and/or for
XX CC inhibiting HIV infection (claimed). The compound that
XX CC the C-type lectin is chosen from mannose, fucose, plant
XX CC antibiotics, sugars, proteins and antibodies against C-
XX SS
SS Query Match
SS Best Local Similarity 100.0%; Score 111; DB 21; Length
SS Matches 111; Conservative 0; Mismatches 0; Indels
OY 449 ATCTACCGAAGCTGACCCAGCTTAAGCTGAGTGGCTCTGAGAAATCCAG 508
DB 229 ATCTACCGAAGCTGACCCAGCTTAAGCTGAGTGGCTCTGAGAAATCCAG 268
OY 509 CTGCAGGAGATCTACGAGTACCTGAGTGAAGCTGCGAGTGGGTGAG 559
DB 289 CTGCAGGAGATCTACGAGTACCTGAGTGAAGCTGCGAGTGGGTGAG 339

```

RESULT 25
AAC65383
ID AAC65383 standard; cDNA; 1215 BP.
XX AAC65383;
AC 12-FEB-2001 (first entry)
DT
XX
DE Human C-type lectin receptor DC-SIGN nucleotide sequence.
KM Human; immunomodulatory; anti-allergic; anti-HIV; anti-
KM antidiabetic; antithyroid; anti-inflammatory; antiarthritic
KM C-type lectin receptor; immunotherapy; immunosuppressive
KM transplant rejection; autoimmune diseases; thyroiditis;
KM rheumatoid arthritis; multiple sclerosis; autoimmune dis-
KM systemic lupus erythematosus; HIV infection; allergy; l-
XX
OS Homo sapiens.
PN W0200063251-A1.
XX
PO 26-OCT-2000.
XX
PF 19-APR-2000; 2000WO-NL00254.
XX
PR 19-APR-1999; 99EP-0201204.

20-JAN-2000; 200005-0176924.

XX XX
PA (OYNI-) UNIV NIJMEGEN.
P1 Fliodor GJ, Geijtenbeek TRM, Van Kooyk Y, Tofenson R.
XX WPI: 2001-6056424/53.
XX P-PSDB; AAB28614.
XX
XX Methods for modulating dendritic cell and T cell interaction by using a
XX compound, such as mannose carbohydrates, that binds to a C-type lectin
XX on a dendritic cell surface, useful in the treatment of autoimmune
XX diseases and allergy -
XX
XX Claim 11: Fig 9; 60pp; English.
XX
XX The present sequence encodes DC-STG1, a 44kDa C-type lectin receptor. N-w
XX methods for modulating dendritic cell and T cell interaction are
XX disclosed. The methods involve using a compound that binds to a C-type
XX lectin on the surface of a dendritic cell. The methods are useful for
XX preventing or inhibiting immune responses to specific antigens, for
XX inducing tolerance, for immunotherapy, for immunosuppression (i.e. for
XX preventing transplant rejection), for the treatment of autoimmune
XX diseases (e.g. thyroiditis, rheumatoid arthritis, multiple sclerosis,
XX autoimmune diabetes, systemic lupus erythematosus), HIV infection and/or
XX for the treatment of allergy. Antibodies are useful for detecting the
XX presence of dendritic cells in a biological sample and for determining
XX the presence and/or expression of the C-type lectins or their fragments
XX or epitopes in a biological sample.
XX
XX Sequence 1217 bp; 315 A; 321 C; 343 T; 290 G; 0 other.
XX
XX Query Match 7.28; Score 111; DB 21; Length 1215;
XX Best Local Similarity 100.0%; Pred No. 1 be-44;
XX Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

449 ATCTACAGAACTGACCTGATTAATGCTGACGCTGAGCTGCTGAGTAAAGTAAAG 508
|||||
229 ATCTACAGAACTGACCTGATTAATGCTGACGCTGAGCTGCTGAGTAAAGTAAAG 288
|||||

509 CTGACGAGGATCTGACCTGAGGACCTGAGCTGAGGCTGAGTAAAGTAAAG 559
|||||
289 CTGACGAGGATCTGACCTGAGGACCTGAGCTGAGGCTGAGTAAAGTAAAG 339
|||||

RESULT 26
AAQ36648
11- AAQ36648 standard; cDNA; 1312 bp.
XX
XX AAQ36648;
XX
XX 08-JUN-1993 (first entry)
XX
XX Sequence encoding non-CD4 glycoprotein gp120 receptor protein.
XX
XX Surface cell protein; glycoprotein receptor protein; gp120 receptor;
XX HIV; diagnosis; therapy; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 42..1256
XX /*tag= a
XX
XX W09301820-A.
XX
XX 04-FEB-1993.
XX
XX 16-JUL-1992; 92WO-US05985.
XX
XX 16-JUL-1991; 91US-0731214.
XX
XX (HRIM) BRISTOL-MYERS SQUIBB CO.

XX
XX Curtis RM;
XX
XX WPI: 1993-058515/07.
XX P-PSDB; AAB32188.
XX
XX
XX Inhibition of non-CD4 mediated HIV infection - dependent on novel
XX gp 120 receptor, for diagnosing and treating HIV infection
XX
XX Disclosure, Figure 3A, 44pp, English.
XX
XX The gp120 receptor protein is present on placental, muscle, neural,
XX brain, dendritic or mucosal cells. It may be produced recombinantly
XX by cloning cDNA isolated from a library of recombinant placental
XX genes. Binding of gp120 to its receptor is inhibited by specific
XX carbohydrates, plant lectins and specific antibodies. It corresp.
XX to a non-CD4 mammalian cell surface protein, contg. ca. 400 AAs and
XX having a mol. wt. of ca. 45 kD, and a binding affinity for gp 120 of
XX kd = 1.32-2 nM.
XX
XX Sequence 1512 bp; 530 A; 548 C; 572 G; 256 T; 0 other.
XX
XX Query Match 7.28; Score 111; DB 14; Length 1312;
XX Best Local Similarity 100.0%; Pred No. 1 be-44;
XX Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

449 ATCTACAGAACTGACCTGATTAATGCTGACGCTGAGCTGCTGAGTAAAGTAAAG 508
|||||
270 ATCTACAGAACTGACCTGATTAATGCTGACGCTGAGCTGCTGAGTAAAGTAAAG 329
|||||

330 CTGACGAGGATCTGACCTGAGGACCTGAGCTGAGGCTGAGTAAAGTAAAG 380
|||||

RESULT 27
AA165461
1D AA165461 standard; DNA; 1312 bp.
XX
XX AA165461;
XX
XX 10-DEC-2001 (first entry)
XX
XX
XX DNA encoding human DC-STG1, a dendritic cell-specific C-type lectin.
XX
XX Human; receptor; DC-STG1; dendritic cell; T lymphocyte; HIV;
XX gp120; C-type lectin; ICAM3; HIV entry; T cell; macrophage;
XX HIV infection; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 42..1256
XX /*tag= a
XX /product= "DC-STG1"
XX
XX W0290164752-A2.
XX
XX 07-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US06322.
XX
XX 02-MAR-2000; 2000US-0517605.
XX
XX (OYNI-) UNIV NEW YORK STATE.
XX (OYNI-) UNIV NIJMEGEN.
XX
XX Litman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;
XX
XX WPI: 2001-602565/68.
XX P-PSDB; AAG79086.
XX
XX An antibody for the treatment or prevention of HIV infection comprises

XX The present invention describes oligonucleotide libraries for detecting
 CC messenger RNAs that populate a (sub)transcriptome, where the
 CC (sub)transcriptome comprises messenger RNAs transcribed from multiple
 CC transcription units that populate a genome. The library comprises
 CC several oligonucleotides, each capable of hybridizing selectively to a
 CC set of messenger RNAs transcribed from a given transcription unit of
 CC the genome, which encodes one of many messenger RNA splice variants.
 CC The oligonucleotide libraries are useful for detecting mRNAs from a
 CC biological sample, in expression profiling studies, in qualitatively or
 CC quantitatively characterizing the corresponding transcriptome, and in
 CC detecting RNA transcripts and splice variants of human or animal
 CC transcriptomes. The libraries may also be used as specialised mini
 CC libraries to detect transcripts of a sub-transcriptome under a
 CC particular biological or pathological state, and so allowing the
 CC detection of tissue- and pathology-specific genes such as those genes
 CC only expressed in specific tissue under a specific pathological
 CC condition; to detect developmental specific genes; and to detect RNA
 CC transcripts and splice variants of a transcriptome of a patient suffering
 CC from a particular disorder. ABN27253 to ABN29589 represent
 CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPAC
 CC at http://wipo.int/pub/published_pat_sequences
 CC
 CC Sequence 60 BP; 11 A; 17 C; 14 G; 18 T; 0 other;

Query Match 3 98; Score 60; DB 24; Length 60;
 Host Local Similarity 100.0%; Prev No 2 9e-19;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1223 GCAAGCTGATCAAGACAAATATTTCTTCTTCTTAAAGCTGATGTTCTAT 1282
 DB 1 GCAGCTGCTTCAATACAAATATTTCTTCTTAAAGCTGATGTTCTAT 60

RESULT 30
 AAH12684/c
 ID AAH12684 standard; cDNA: 597 BP.

XX AAH12684;

XX 26-JUN-2001 (first entry)

XX Human cDNA clone (3'-primer) SEQ ID NO:9519.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EPI074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP 0114126

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELIX) HELIX RES INST.

XX Ota I, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI: 2001-318749/34.

Primer sets for synthesising polynucleotides, particularly the 5602
 full-length cDNAs defined in the specification, and for the detection
 and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -
 XX Claim 3; SEQ ID 9519; 2537bp - CD ROM, English.

XX The present invention describes primer sets for synthesising 5602
 CC full-length cDNAs defined in the specification where a primer set
 CC comprises: (a) an oligo-p primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5' end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3' end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5' end sequence/3' end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH1076 to AAH1478 and
 CC AAH1484 to AAH18742 represent human cDNA sequences; AAB2446 to
 CC AAB35893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides all of which are used in the exemplification
 CC of the present invention.

XX Sequence 597 BP; 186 A; 147 C; 118 G; 128 T; 18 other;

Query Match 2 89; Score 43; DB 22; Length 597;
 Host Local Similarity 100.0%; Prev No 7 4e-11;
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1491 TGAGTGTACCTTTATATTAAGTGTAAAGTTGTAAGTCG 1533
 DB 43 TGAGTGTACCTTTATATTAAGTGTAAAGTTGTAAGTCG 1

RESULT 31
 AAI84994/c
 ID AAI84994 standard; cDNA: 338 BP.

XX AAI84994;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 5054.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; hematopoiesis;

XX tissue growth factor; immunomodulatory; cancer; leukemia;

XX nervous system disorders; arthritis; inflammation; ss.

XX Homo sapiens.

XX W0200064825 A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-0504927.

XX 28-FEB-2000; 2000US-0515126.

XX 18-MAY-2000; 2000US-0577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-514838/56.

XX P-PDB: AAB05063.

Isolated nucleic acids and polypeptides, useful for preventing
 diagnosis and treating e.g. leukemia, inflammation and immune

disorders .

XX Claim 1, SEQ ID NO 5054: 1399bp . Sequence listing: English.

XX The invention relates to human polypeptides (AA179941-AA193841) and

CC the encoded proteins (AA0001-AA019410) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polypeptides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, hematopoietic regulatory

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activity/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp://patent/pub/published_pat_sequences.

XX Sequence 338 bp: 1st A, 65 C, 102 G, 68 T, 2 other.

XX Query Match: 2 59 Score 39, 2R 22, Length 338;

XX Best Local Similarity: 100.0%, Freq No: 7, 1e-09;

XX Matches: 39; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

UY 1355 TTCTAGCTGACAGCAGCAGCAATTAATTCCTCCATTC 1393

DB 256 TTCTAGCTGACAGCAGCAGCAATTAATTCCTCCATTC 218

RESULT 32

AAS36709/c

ID AAS36709 standard. ENA, 32190 BP.

AC AAS36709.

DI 17-DEC-2001 (first entry)

DE Human cardiovascular system antigen genomic DNA SEQ ID NO 2209.

XX

KM Cardiovascular system antigen: human, mouse, rabbit, goat; horse; cat;

KM chicken, sheep, immunosuppressive, antiarthritis; vasotrophic; dog;

KM antineumatic; antiproliferative; cytostatic; cardiact; neuroprotective;

KM cerebroprotective; neurotrophic; antibacterial; virocid; fungicide; cancer;

KM ophthalmological; vulnecary; gene therapy; autoimmune disease; neoplasm;

KM hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;

KM cerebrovascular disorder; nervous system disorder; bacterial infection;

KM fungal infection; viral infection; ocular disorder; endocrine disorder;

KM gastrointestinal disorder; renal disorder; respiratory disorder;

KM wound healing; skin aging; organ transplantation; tissue regeneration;

KM anti-infertility.

XX

OS Homo sapiens

XX

PN WO200155321-A2.

XX

PD 02-AUG-2001

XX

PE 17-JAN-2001, 2001W1-0801340

XX

XX 31-JAN-2000; 2000US-0179565

PR 04-FEB-2000; 2000US-0180628

PR 24-FEB-2000; 2000US-0184664

PR 02-MAR-2000; 2000US-0186750

PR 16-MAR-2000; 2000US-0189874

PR 17-MAR-2000; 2000US-0190076

PR 18-APR-2000; 2000US-0198123

PR 19-MAY-2000; 2000US-0205515

PR 07-JUN-2000; 2000US-0209467

PR 28-JUN-2000; 2000US-0214886

PR 30-JUN-2000; 2000US-0215135

PR 07-JUL-2000; 2000US-0215647

PR 07-JUL-2000; 2000US-0215880

PR 11-JUL-2000; 2000US-0217487

PR 11-JUL-2000; 2000US-0217496

PR 14-JUL-2000; 2000US-0218290

PR 26-JUL-2000; 2000US-0220663

PR 26-JUL-2000; 2000US-0220964

PR 14-AUG-2000; 2000US-0224518

PR 14-AUG-2000; 2000US-0224519

PR 14-AUG-2000; 2000US-0225213

PR 14-AUG-2000; 2000US-0225414

PR 14-AUG-2000; 2000US-0225266

PR 14-AUG-2000; 2000US-0225267

PR 14-AUG-2000; 2000US-0225468

PR 14-AUG-2000; 2000US-0225270

PR 14-AUG-2000; 2000US-0225477

PR 14-AUG-2000; 2000US-0225477

PR 14-AUG-2000; 2000US-0225758

PR 14-AUG-2000; 2000US-0225759

PR 18-AUG-2000; 2000US-0226279

PR 22-AUG-2000; 2000US-0226681

PR 22-AUG-2000; 2000US-0226688

PR 22-AUG-2000; 2000US-0227182

PR 23-AUG-2000; 2000US-0227009

PR 30-AUG-2000; 2000US-0228424

PR 01-SEP-2000; 2000US-0229287

PR 01-SEP-2000; 2000US-0229343

PR 01-SEP-2000; 2000US-0229344

PR 01-SEP-2000; 2000US-0229345

PR 01-SEP-2000; 2000US-0229509

PR 05-SEP-2000; 2000US-0229513

PR 05-SEP-2000; 2000US-0230437

PR 06-SEP-2000; 2000US-0230438

PR 06-SEP-2000; 2000US-0231242

PR 08-SEP-2000; 2000US-0231243

PR 08-SEP-2000; 2000US-0241244

PR 08-SEP-2000; 2000US-0231413

PR 08-SEP-2000; 2000US-0231414

PR 08-SEP-2000; 2000US-0232080

PR 08-SEP-2000; 2000US-0232081

PR 12-SEP-2000; 2000US-0231968

PR 14-SEP-2000; 2000US-0233497

PR 14-SEP-2000; 2000US-0233498

PR 14-SEP-2000; 2000US-0233399

PR 14-SEP-2000; 2000US-0233400

PR 14-SEP-2000; 2000US-0233401

PR 14-SEP-2000; 2000US-0233863

PR 14-SEP-2000; 2000US-0233864

PR 14-SEP-2000; 2000US-0233865

PR 21-SEP-2000; 2000US-0234223

PR 21-SEP-2000; 2000US-0234274

PR 25-SEP-2000; 2000US-0234997

PR 25-SEP-2000; 2000US-0234998

PR 25-SEP-2000; 2000US-0235484

PR 26-SEP-2000; 2000US-0235834

PR 27-SEP-2000; 2000US-0235836

PR 27-SEP-2000; 2000US-0235837

PR 29-SEP-2000; 2000US-0236367

PR 29-SEP-2000; 2000US-0236368

PR 29-SEP-2000; 2000US-0236369

PR 29-SEP-2000; 2000US-0236370

PR 02-OCT-2000; 2000US-0236802

PR 02-OCT-2000; 2000US-0237037

PR 02-OCT-2000; 2000US-0237038

PR 02-OCT-2000; 2000US-0237039

PR 02-OCT-2000; 2000US-0237040

PR 13-OCT-2000; 2000US-0237945

PR 13-OCT-2000; 2000US-0237947

PR 20-OCT-2000; 2000US-0240560

PR 20-OCT-2000; 2000US-0241785

PR 20-OCT-2000; 2000US-0241786

PR 20-OCT-2000; 2000US-0241787

PR 20-OCT-2000; 2000US-0241808

PR 20-OCT-2000; 2000US-0241809

20-CT-2000: 2000US-0241826.
 EP 21-NOV-2000: 2000US-0244617.
 EP 08-NOV-2000: 2000US-0246474.
 EP 08-NOV-2000: 2000US-0246475.
 ER 08-NOV-2000: 2000US-0246476.
 ER 08-NOV-2000: 2000US-0246477.
 ER 08-NOV-2000: 2000US-0246478.
 ER 08-NOV-2000: 2000US-0246523.
 ER 08-NOV-2000: 2000US-0246524.
 ER 08-NOV-2000: 2000US-0246525.
 ER 08-NOV-2000: 2000US-0246526.
 ER 08-NOV-2000: 2000US-0246527.
 ER 08-NOV-2000: 2000US-0246528.
 ER 08-NOV-2000: 2000US-0246532.
 ER 08-NOV-2000: 2000US-0246609.
 ER 08-NOV-2000: 2000US-0246610.
 ER 08-NOV-2000: 2000US-0246611.
 ER 08-NOV-2000: 2000US-0246613.
 ER 17-NOV-2000: 2000US-0249207.
 ER 17-NOV-2000: 2000US-0249208.
 ER 17-NOV-2000: 2000US-0249209.
 ER 17-NOV-2000: 2000US-0249210.
 ER 17-NOV-2000: 2000US-0249211.
 ER 17-NOV-2000: 2000US-0249212.
 ER 17-NOV-2000: 2000US-0249213.
 ER 17-NOV-2000: 2000US-0249214.
 ER 17-NOV-2000: 2000US-0249215.
 ER 17-NOV-2000: 2000US-0249216.
 ER 17-NOV-2000: 2000US-0249217.
 ER 17-NOV-2000: 2000US-0249218.
 ER 17-NOV-2000: 2000US-0249244.
 ER 17-NOV-2000: 2000US-0249245.
 ER 17-NOV-2000: 2000US-0249264.
 ER 17-NOV-2000: 2000US-0249265.
 ER 17-NOV-2000: 2000US-0249297.
 ER 17-NOV-2000: 2000US-0249299.
 ER 17-NOV-2000: 2000US-0249300.
 ER 01-DEC-2000: 2000US-0250160.
 ER 01-DEC-2000: 2000US-0250391.
 ER 05-DEC-2000: 2000US-0251040.
 ER 05-DEC-2000: 2000US-0251988.
 ER 05-DEC-2000: 2000US-0256719.
 ER 06-DEC-2000: 2000US-0251479.
 ER 08-DEC-2000: 2000US-0251456.
 ER 08-DEC-2000: 2000US-0251868.
 ER 08-DEC-2000: 2000US-0251869.
 ER 08-DEC-2000: 2000US-0251989.
 ER 08-DEC-2000: 2000US-0251990.
 ER 11-DEC-2000: 2000US-0254097.
 ER 05-JAN-2001: 2001US-0254678.
 XX
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Raben SM.
 XX
 DE WPI; 2001-451930/48.
 XX
 XX New cardiovascular system related polynucleotides and polypeptides,
 PT useful for diagnosing, treating and/or preventing disorders of the
 PT cardiovascular system -
 XX
 PS Claim 1: SEQ ID NO 2209; 674bp; English.
 XX
 CC Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode
 CC the cardiovascular system antigen polypeptides of the invention.
 CC Cardiovascular system antigens and their associated polynucleotides are
 CC useful in the diagnosis, treatment and prevention of various types of
 CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs.
 CC Chickens or sheep. A pathological condition can be determined by
 CC detecting the presence or absence of a mutation in a cardiovascular
 CC system antigen polynucleotide. The treatable disorders include autoimmune
 CC diseases such as rheumatoid arthritis, hyperproliferative disorders such as
 CC as neoplasms of the breast or liver, cardiovascular disorders such as

CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,
 CC nervous system disorders such as Alzheimer's disease, infections caused
 CC by bacteria, viruses and fungi, cardiac disorders such as normal
 CC infection, endocrine disorders such as premature labour and infertility,
 CC gastrointestinal disorders such as Crohn's disease, renal disorders such
 CC as glomerulonephritis and respiratory disorders such as asthma and
 CC fluency. The polypeptides can also be used to aid wound healing, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, to regenerate tissues and in chemotaxis.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pcl_sequences.
 XX
 Query Match 1.8%, Score 28, DB 23; Length 32190;
 Best local similarity 100.0%, T-vec No. 0.00019;
 Matches 28, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
 QY 1472 TCTTCATTTGAGCTGTTCTGATTTGA 1499
 ID 10561 TCTTCATTTGAGCTGTTCTGATTTGA 10334
 RESULT 33
 ABL84083/c
 ID ABL84083 standard, cDNA, 472 bp.
 XX
 AC ABL84083;
 XX
 ST 17 NOV 2002 (first entry)
 XX
 EE Human ovarian cancer related cDNA clone SEQ ID NO:7061.
 XX
 KW Human, ovarian cancer, ovarian tumour, cytostatic, gene, ss.
 XX
 OS Homo sapiens.
 XX
 FN WC200192581 A2.
 XX
 PD 06 DEC 2001.
 XX
 FT 29 MAY 2001; 2001US 0317756.
 XX
 FE 26 MAY 2002; 2002US-207484P.
 XX
 PA (CORI-) CORIAXA CORP.
 XX
 PI Algate PA, Harlocker SL, Jones R;
 XX
 DE WPI; 2002-022075/16.
 XX
 XX Composition for therapy and diagnosis of ovarian cancer comprising
 PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 PT polypeptide, and body specific to polypeptide or 1 cell expressing
 PT polypeptide -
 XX
 PS Claim 1: SEQ ID 7061; 489bp; English.
 XX
 CC The present invention describes a composition (I) comprising: carriers
 CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
 CC polynucleotide encoded by a polynucleotide (III) having a cDNA sequence
 CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
 CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
 CC population of (II), or antigen presenting cells that express (II).
 CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
 CC (S1) can be used for detecting ovarian cancer in a patient's biological
 CC sample preferably serum or ovarian tissue. The method comprises
 CC contacting a biological sample from a patient with (IV), detecting the
 CC amount of polynucleotide hybridising to (IV) and comparing the amount to
 CC a predetermined cutoff value and thereby detecting ovarian cancer in the
 CC patient, where the amount of polynucleotide hybridising to (IV) is
 CC detected preferably by polymerase chain reaction (PCR). (I) comprising
 CC (III) and/or (II) is useful for stimulating and/or expanding T cells

CC specific for an ovarian tumour protein comprising contacting T cells
 CC with (II) or (II) is useful in design and preparation of
 CC ribozyme molecules for inhibiting expression of the tumour polypeptides
 CC and proteins in tumour cells; and to isolate a full length gene from a
 CC suitable library e.g., a tumour cDNA library using well known
 CC techniques.
 XX
 XX Sequence 472 BP, 124 A, 110 C, 93 G, 145 T, 0 other.
 XX
 Query Match 1.78; Score 26; FR 24; Length 472;
 Best Local Similarity 100.0%; Pred No 0.019;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1474 TTCATTTGGCTGTTCTGACTTCTGA 1499
 DB 219 TTGCAATTCCTGCTTCTGACTTCTGA 194
 XX
 RESULT 14
 AAD03249
 ID AAD03249 standard; cDNA: 59 BP.
 AC AAD03249;
 XX
 DT 13-JUN-2001 (first entry)
 XX
 DE Mouse type II integral membrane protein, CIRE-like fragment #22 cDNA.
 XX
 KW Mouse; type II integral membrane protein; CIRE, anti-allergic vaccine;
 KW immunosuppressive; autoimmune disorder; allergy; viral infection; APC;
 KW antigen presenting cell; bacterial infection; macrophage; dendritic cell;
 KW DC; ss.
 XX
 OS Mus musculus.
 XX
 PN W0200119869-A1.
 XX
 PD 22-MAR-2001.
 XX
 PF 13-SEP-2000; 2000WO-A001096.
 XX
 PR 13-SEP-1999; 99AU-0002788.
 XX
 PA (COON-) COUNCIL QUEENSLAND INST MEDICAL RES.
 XX
 P1 Caminschi I, Wright MD, Shortman KD;
 XX
 DK WPI: 2001-244788/25.
 XX
 PT Novel type II integral membrane protein, CIRE, preferentially expressed
 PT in dendritic cells, macrophages and their precursors stimulatory to T
 PT cells, useful for identifying compounds having immunoregulatory
 PT activity -
 XX
 PS Disclosure; Page 28; 53pp; English.
 XX
 CC The present cDNA sequence encodes mouse CIRE-like fragment derived
 CC from a protein expressed on dendritic cells which has similar function
 CC to CIRE protein.
 CC
 CC The invention relates to mouse type II integral membrane protein, CIRE
 CC and its corresponding cDNA molecule which is preferentially expressed in
 CC dendritic cells, macrophages and their precursors stimulatory to T cells.
 CC CIRE is useful for screening immunoregulatory compounds and as a marker
 CC for dendritic cell (DC) population. CIRE specific ligand such as a
 CC monoclonal antibodies are useful for isolating an antigen presenting cell
 CC (APC) from a biological sample. CIRE DNA is also used as tools to analyse
 CC the properties and functions of the CIRE gene/protein. For example, the
 CC DNA molecules are used to generate animal models, which lack functional
 CC CIRE genes and to isolate regulatory regions of the CIRE gene. Such
 CC regulatory regions are used to selectively express exogenous genes in DC
 CC or APC. The invention also relates to methods for modulating an immune
 CC response in a subject by binding and inhibiting the function of an APC
 CC such as a myeloid DC. The CIRE-specific ligands are useful for modulating

CC immune response by interfering with the function, migration
 CC or DC or APC. The ligands bind to and interfere with the
 CC myeloid dendritic cells such that antigen processing is
 CC lymphoid DC which leads to immune suppression and allergy;
 CC allergies and autoimmune disorders. The ligands are also
 CC or enhance the interaction of viruses or bacteria with the
 CC resistance to infection. These ligands are also used to
 CC such as vaccine components, to DC or APC.
 XX
 XX Sequence 55 BP, 3 A, 20 C, 13 G, 17 T, 0 other;
 XX
 Query Match 1.58; Score 23; FR 22; Length 54;
 Best Local Similarity 100.0%; Pred No 0.6;
 Matches 23; Conservative 0; Mismatches 0; Indels 0;
 XX
 QY 867 TCTTCCAGCAAGCACTGTTACTTC 879
 DB 32 TCTTCCAGCAAGCACTGTTACTTC 54
 XX
 RESULT 35
 AAD03228
 ID AAD03228 standard; cDNA: 154 BP.
 AC AAD03228;
 XX
 DT 13-JUN-2001 (first entry)
 XX
 DE Mouse CIRE like fragment #1 encoding cDNA clone, RP23-458N1.
 XX
 KW Mouse; type II integral membrane protein; CIRE, anti-allergic
 KW immunosuppressive; autoimmune disorder; allergy; viral
 KW antigen presenting cell; bacterial infection; macrophage
 KW DC; ss.
 XX
 OS Mus musculus.
 XX
 PN W0200119869-A1.
 XX
 PD 22-MAR-2001.
 XX
 PF 13-SEP-2000; 2000WO-A001096.
 XX
 PR 13-SEP-1999; 99AU-0002788.
 XX
 PA (COON-) COUNCIL QUEENSLAND INST MEDICAL RES.
 XX
 P1 Caminschi I, Wright MD, Shortman KD;
 XX
 DK WPI: 2001-244788/25.
 XX
 PT Novel type II integral membrane protein, CIRE, preferentially expressed
 PT in dendritic cells, macrophages and their precursors stimulatory to T
 PT cells, useful for identifying compounds having immunoregulatory
 PT activity -
 XX
 PS Disclosure; Page 24; 53pp; English.
 XX
 CC The present cDNA sequence encodes CIRE-like fragment 11
 CC RP23-458N1. This fragment is derived from a protein expressed
 CC dendritic cells which has similar function to CIRE protein.
 CC The invention relates to mouse type II integral membrane
 CC and its corresponding cDNA molecule which is preferentially expressed in
 CC dendritic cells, macrophages and their precursors stimulatory to T cells.
 CC CIRE is useful for screening immunoregulatory compounds

for dendritic cell (DC) population. CIRE specific ligand such as monoclonal antibodies are useful for isolating an antigen presenting cell (APC) from a biological sample. CIRE DNA is also used as tools to analyse the properties and functions of the CIRE gene/protein. For example, the CIRE molecules are used to generate animal models, which lack functional CIRE genes and to isolate regulatory regions of the CIRE gene. Such regulatory regions are used to selectively express exogenous genes in DC or APC. The invention also relates to methods for modulating an immune response in a subject by binding and inhibiting the function of an APC such as a myeloid DC. The CIRE-specific ligands are useful for modulating immune response by interfering with the function, migration or maturation of DC or APC. The ligands bind to and interfere with the function of myeloid dendritic cells such that antigen processing is undertaken by lymphoid DC which leads to immune suppression and anergy for treating allergies and autoimmune disorders. The ligands are also useful to block or enhance the interaction of viruses or bacteria with DC, for improving resistance to infection. These ligands are also used to target molecules such as vaccine components, to DC or APC.

Query Match 1.5% Score 23, DB 23, Length 153;
Best Local Similarity 100.0%, Prod No. 0.59;
Matches 23, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

867 TCTTCAGCAAGCAACTGTACTTC 889
|||||
38 TCTTCAGCAAGCAACTGTACTTC 60

RESULT 36
AAS80491/c
ID AAS80491 standard; cDNA: 2879 BP.

AC AAS80491;

DT 11-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #16295.

KW Human, chromosome mapping; gene mapping; gene therapy; forensic;

EV food supplement, medical imaging, diagnostic, genetic disorder, ss

XX Homo sapiens.

PN W0200175067-A2.

XX 11-OCT-2001.

PE 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217

PR 23-AUG-2000; 2000US-0649167.

XX (HYSK-) HYSKO INC.

PI Dmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DE P-PSDB: ABG16394.

XX New isolated polynucleotide and encoded polypeptides, useful in PT diagnostics, forensics, gene mapping, identification of mutations PT responsible for genetic disorders or other traits and to assess PT biodiversity

XX Claim 1; SEQ ID No 16295, 103pp, English.

XX The invention relates to isolated polynucleotide (I) and CC polypeptide (II) sequences. (I) is useful as hybridisation probes, CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome CC and gene mapping, and in recombinant production of (II). The CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques CC to restore normal activity of (II) or to treat disease states involving CC (II). (II) is useful for generating antibodies against it, detecting or CC quantitating a polypeptide in tissue, as molecular weight markers and as CC a food supplement. (II) and its binding partners are useful in medical CC imaging of sites expressing (II). (I) and (II) are useful for treating CC disorders involving aberrant protein expression or biological activity. CC The polypeptide and polynucleotide sequences have applications in CC diagnostics, forensics, gene mapping, identification of mutations CC responsible for genetic disorders or other traits to assess biodiversity CC and to produce other types of data and products dependent on DNA and CC amino acid sequences. AAS64197-AAS94564 represent novel human CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed CC specification, but was obtained in electronic format directly from WIPO CC at the wipo.int/patlib/checked_pat_sequences.

XX Sequence 2879 BP, A: 279 C: 842 G: 517 T: 0 other;

Query Match 1.1%, Score 22, DB 23, Length 2879;
Best Local Similarity 100.0%, Prod No. 1.8;
Matches 22, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

867 TCTTCAGCAAGCAACTGTACTTC 1493
|||||
2211 TCTTCAGCAAGCAACTGTACTTC 2190

RESULT 37
ABK34644/c
ID ABK34644 standard; cDNA: 2461 BP.

AC ABK34644;

DT 08-MAY-2002 (first entry)

DE Human cDNA for novel secreted protein, SEQ ID 413.

XX Human (ss) gene; secreted protein; immune deficiency; viral infection;

KW bacterial infection, fungal infection, autoimmune disorder, burn;

KW rheumatoid arthritis, multiple sclerosis, autoimmune thyroiditis;

KW diabetes, allergy, asthma, neurodegenerative disease, chronic disease;

KW Alzheimer's disease, Parkinson's disease, liver fibrosis, tumour;

KW coagulation disorder, haemophilia, inflammatory disorder, ulcer;

KW tissue regeneration; wound healing, hematopoiesis; myeloid deficiency;

XX Lymphoid cell deficiency.

XX Homo sapiens.

PN W0200177290 A2.

XX 18-OCT-2001.

PE 29-MAR-2001; 2001WO-US10295.

XX 06-APR-2000; 2000US-194941P.

XX (GENY) GENETICS INST INC.

PI Wong CG, Clark RF, Fochtel K, Agostino MJ, Howes SH, Resnik RJ;

XX Gulakota K, Graham JR;

XX WPI; 2002-179422/23.

XX Six hundred and twenty five polynucleotides derived from a variety of PT human tissue sources which encode secreted proteins, useful for PT treating immune deficiencies and disorders such as autoimmune disorders PT

XX Claim 1; Page 213; 339pp; English.
XX The invention relates to 625 polynucleotides which have been derived from CC a variety of human tissue sources and which encode novel secreted

CC proteins, their complements and sequences that hybridise to them.
 CC Also included are a vector comprising the polynucleotide, a host cell
 CC transformed with the vector, the proteins encoded by the
 CC polynucleotides, antibodies that bind to the proteins and identification
 CC of modulators of the proteins or the expression of the polynucleotide.
 CC The polynucleotides can be used as probes for the identification
 CC and isolation of full length cDNA and genomic DNA. The polynucleotides
 CC and proteins can also be used as nutritional supplements. The protein
 CC is useful in the treatment of various immune deficiencies and disorders
 CC such as viral infections, bacterial infections, fungal infections,
 CC autoimmune disorders (e.g. rheumatoid arthritis, multiple sclerosis,
 CC autoimmune thyroiditis and diabetes) and allergic reactions and
 CC conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment
 CC of burns, incisions and ulcers. The proteins are also useful for
 CC regulating haematopoiesis, for treating myeloid or lymphoid cell
 CC deficiencies. The present sequence is one of the 625 cDNA sequences
 CC encoding a secreted protein.

XX Sequence 2361 BP; 637 A; 535 C; 491 G; 697 T; 1 other;

Query Match 1.48; Score 21; DB 24; Length 2361.
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1479 TTGGCTGTTCTGAGTTGTA 1499
 |||
 DB 852 TTGGCTGTTCTGAGTTGTA 832

RESULT 38

AAC30751/c
 ID AAC30751 standard; cDNA; 157 BP.

AC AAC30751;

XX 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 34826.

KW Human; 5' EST; expressed sequence tag, secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping. ss.

XX Homo sapiens.

FN EP103401-A2.

XX 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 990US 0122487.

PA (GENSET) GENSET.

PI Dumas Milne Edwards J. Buchert A. Giordano J.;

DE WPI: 2000-503381/45.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping purposes -

XX Claim 1: SEQ ID 34826; 71bp + CD-ROM. English

XX The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or poly(A) RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTP)

CC of the mRNA because they are often obtained from oligo-c
 CC libraries. Such ESTs are not well suited for isolating c
 CC derived from the 5' ends of mRNAs and even in those cases
 CC cDNA sequences have been obtained, the full 5' UTR is r
 CC 5' ESTs are derived from mRNAs with intact 5' ends and ca
 CC used to obtain full length cDNAs and genomic DNAs. 5' EST
 CC in diagnostic, forensic, gene therapy and chromosome map
 CC They are used to obtain upstream regulatory sequences and
 CC expression and secretion vectors.

XX Sequence 157 BP; 21 A; 62 C; 43 G; 31 T; 0 other;

Query Match 1.38; Score 20; DB 21; Length
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 20; Conservative 0; Mismatches 0; Indels 0;

QY 603 CCGGGCTGAAGCTGCTGACTG 622
 |||
 DB 26 CCGGGCTGAAGCTGCTGACTG 7

RESULT 39

AAK59336/c
 ID AAK59336 standard; cDNA; 415 BP.

AC AAK59336;

XX 06-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen encoding cDNA SEQ 1

XX Human; immune; haematopoietic; immune/haematopoietic ant

KW cytosolic; gene therapy; vaccines; metastasis; ss.

OS Homo sapiens.

PN W0200157182-A2.

PD 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01364.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198124.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 01-JUN-2000; 2000US-0215145.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218240.

PR 26-JUL-2000; 2000US-0220047.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

RESULT 40
AAK75081/C
ID AAK75081 standard; DNA: 319 RP
XX AAK75081,
XX 07-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen gene-wide sequence. SEQ ID No: 29833
XX
XX Human: immune, haematopoietic; immune/haematopoietic antigen: cancer;
KM cytostatic, gene therapy, vaccine; metastasis; ds
XX
XX Homo sapiens.
XX
XX W0208157182.A2.
XX
XX 09-AUG-2001
XX
XX 17-JAN-2001; 2001W0-MS01454
XX
XX 41-JAN-2000; 2000US-0179665
XX 04-FEB-2000; 2000US-0180628
XX 24-FEB-2000; 2000US-0184664
XX 02-MAR-2000; 2000US-0186360
XX 16-MAR-2000; 2000US-0189874
XX 17-MAR-2000; 2000US-0190076
XX 18-APR-2000; 2000US-0198123
XX 19-MAY-2000; 2000US-0205515
XX 07-JUN-2000; 2000US-0209467
XX 28-JUN-2000; 2000US-0214886
XX 30-JUN-2000; 2000US-0215145
XX 07-JUL-2000; 2000US-0216647
XX 07-JUL-2000; 2000US-0216880
XX 11-JUL-2000; 2000US-0217487
XX 11-JUL-2000; 2000US-0217496
XX 14-JUL-2000; 2000US-0218290
XX 26-JUL-2000; 2000US-0220963
XX 26-JUL-2000; 2000US-0220964
XX 14-AUG-2000; 2000US-0224518
XX 14-AUG-2000; 2000US-0224519
XX 14-AUG-2000; 2000US-0225213
XX 14-AUG-2000; 2000US-0225214
XX 14-AUG-2000; 2000US-0225466
XX 14-AUG-2000; 2000US-0225467
XX 14-AUG-2000; 2000US-0225468
XX 14-AUG-2000; 2000US-0225470
XX 14-AUG-2000; 2000US-0225447
XX 14-AUG-2000; 2000US-0225757
XX 14-AUG-2000; 2000US-0225758
XX 14-AUG-2000; 2000US-0225759
XX 14-AUG-2000; 2000US-0225779
XX 22-AUG-2000; 2000US-0226683
XX 22-AUG-2000; 2000US-0226868
XX 22-AUG-2000; 2000US-0227182
XX 23-AUG-2000; 2000US-0227009
XX 30-AUG-2000; 2000US-0228924
XX 01-SEP-2000; 2000US-0229287
XX 01-SEP-2000; 2000US-0229343
XX 01-SEP-2000; 2000US-0229344
XX 01-SEP-2000; 2000US-0229345
XX 05-SEP-2000; 2000US-0229609
XX 05-SEP-2000; 2000US-0229613
XX 06-SEP-2000; 2000US-0230437
XX 06-SEP-2000; 2000US-0230439
XX 08-SEP-2000; 2000US-0231242
XX 08-SEP-2000; 2000US-0231243
XX 08-SEP-2000; 2000US-0231244
XX 08-SEP-2000; 2000US-0231413
XX 08-SEP-2000; 2000US-0231414
XX 08-SEP-2000; 2000US-0232880

PR 08-SEP-2000; 2000US-0232081
PR 12-SEP-2000; 2000US-0231968
PR 14-SEP-2000; 2000US-0233397
PR 14-SEP-2000; 2000US-0233398
PR 14-SEP-2000; 2000US-0233399
PR 14-SEP-2000; 2000US-0233400
PR 14-SEP-2000; 2000US-0233401
PR 14-SEP-2000; 2000US-0233063
PR 14-SEP-2000; 2000US-0233064
PR 14-SEP-2000; 2000US-0233065
PR 21-SEP-2000; 2000US-0234223
PR 21-SEP-2000; 2000US-0234274
PR 25-SEP-2000; 2000US-0234997
PR 25-SEP-2000; 2000US-0234998
PR 26-SEP-2000; 2000US-0235484
PR 27-SEP-2000; 2000US-0235834
PR 27-SEP-2000; 2000US-0235836
PR 29-SEP-2000; 2000US-0236327
PR 29-SEP-2000; 2000US-0236367
PR 29-SEP-2000; 2000US-0236368
PR 29-SEP-2000; 2000US-0236369
PR 29-SEP-2000; 2000US-0236370
PR 02-OCT-2000; 2000US-0236802
PR 02-OCT-2000; 2000US-0237037
PR 02-OCT-2000; 2000US-0237038
PR 02-OCT-2000; 2000US-0237039
PR 02-OCT-2000; 2000US-0237040
PR 13-OCT-2000; 2000US-0239935
PR 13-OCT-2000; 2000US-0239937
PR 20-OCT-2000; 2000US-0240960
PR 20-OCT-2000; 2000US-0241221
PR 20-OCT-2000; 2000US-0241785
PR 20-OCT-2000; 2000US-0241786
PR 20-OCT-2000; 2000US-0241787
PR 20-OCT-2000; 2000US-0241808
PR 20-OCT-2000; 2000US-0241809
PR 20-OCT-2000; 2000US-0241826
PR 01-NOV-2000; 2000US-0244517
PR 08-NOV-2000; 2000US-0246474
PR 08-NOV-2000; 2000US-0246475
PR 08-NOV-2000; 2000US-0246476
PR 08-NOV-2000; 2000US-0246477
PR 08-NOV-2000; 2000US-0246478
PR 08-NOV-2000; 2000US-0246523
PR 08-NOV-2000; 2000US-0246524
PR 08-NOV-2000; 2000US-0246525
PR 08-NOV-2000; 2000US-0246526
PR 08-NOV-2000; 2000US-0246527
PR 08-NOV-2000; 2000US-0246528
PR 08-NOV-2000; 2000US-0246532
PR 08-NOV-2000; 2000US-0246609
PR 08-NOV-2000; 2000US-0246610
PR 08-NOV-2000; 2000US-0246611
PR 08-NOV-2000; 2000US-0246613
PR 17-NOV-2000; 2000US-0249207
PR 17-NOV-2000; 2000US-0249208
PR 17-NOV-2000; 2000US-0249209
PR 17-NOV-2000; 2000US-0249210
PR 17-NOV-2000; 2000US-0249211
PR 17-NOV-2000; 2000US-0249212
PR 17-NOV-2000; 2000US-0249213
PR 17-NOV-2000; 2000US-0249214
PR 17-NOV-2000; 2000US-0249215
PR 17-NOV-2000; 2000US-0249216
PR 17-NOV-2000; 2000US-0249217
PR 17-NOV-2000; 2000US-0249218
PR 17-NOV-2000; 2000US-0249244
PR 17-NOV-2000; 2000US-0249245
PR 17-NOV-2000; 2000US-0249264
PR 17-NOV-2000; 2000US-0249265
PR 17-NOV-2000; 2000US-0249297
PR 17-NOV-2000; 2000US-0249299
PR 17-NOV-2000; 2000US-0249300


```

PR 29-SEP-2000: 2000US-0236369.
PR 29-SEP-2000: 2000US-0236370.
PR 02-OCT-2000: 2000US-0236802.
PR 02-OCT-2000: 2000US-0237037.
PR 02-OCT-2000: 2000US-0237038.
PR 02-OCT-2000: 2000US-0237039.
PR 02-OCT-2000: 2000US-0237040.
PR 13-OCT-2000: 2000US-0239935.
PR 13-OCT-2000: 2000US-0239937.
PR 20-OCT-2000: 2000US-0240960.
PR 20-OCT-2000: 2000US-0241221.
PR 20-OCT-2000: 2000US-0241785.
PR 20-OCT-2000: 2000US-0241786.
PR 20-OCT-2000: 2000US-0241787.
PR 20-OCT-2000: 2000US-0241808.
PR 20-OCT-2000: 2000US-0241809.
PR 20-OCT-2000: 2000US-0241826.
PR 01-NOV-2000: 2000US-0244617.
PR 08-NOV-2000: 2000US-0244674.
PR 08-NOV-2000: 2000US-0244675.
PR 08-NOV-2000: 2000US-0244676.
PR 08-NOV-2000: 2000US-0244677.
PR 08-NOV-2000: 2000US-0244678.
PR 08-NOV-2000: 2000US-0244623.
PR 08-NOV-2000: 2000US-0244624.
PR 08-NOV-2000: 2000US-0244625.
PR 08-NOV-2000: 2000US-0244626.
PR 08-NOV-2000: 2000US-0244627.
PR 08-NOV-2000: 2000US-0244628.
PR 08-NOV-2000: 2000US-0244632.
PR 08-NOV-2000: 2000US-0244609.
PR 08-NOV-2000: 2000US-0244610.
PR 08-NOV-2000: 2000US-0244611.
PR 08-NOV-2000: 2000US-0244613.
PR 17-NOV-2000: 2000US-0249207.
PR 17-NOV-2000: 2000US-0249208.
PR 17-NOV-2000: 2000US-0249209.
PR 17-NOV-2000: 2000US-0249210.
PR 17-NOV-2000: 2000US-0249211.
PR 17-NOV-2000: 2000US-0249212.
PR 17-NOV-2000: 2000US-0249213.
PR 17-NOV-2000: 2000US-0249214.
PR 17-NOV-2000: 2000US-0249215.
PR 17-NOV-2000: 2000US-0249216.
PR 17-NOV-2000: 2000US-0249217.
PR 17-NOV-2000: 2000US-0249218.
PR 17-NOV-2000: 2000US-0249244.
PR 17-NOV-2000: 2000US-0249245.
PR 17-NOV-2000: 2000US-0249264.
PR 17-NOV-2000: 2000US-0249265.
PR 17-NOV-2000: 2000US-0249267.
PR 17-NOV-2000: 2000US-0249299.
PR 17-NOV-2000: 2000US-0249300.
PR 01-DEC-2000: 2000US-0250160.
PR 01-DEC-2000: 2000US-0250161.
PR 05-DEC-2000: 2000US-0251040.
PR 05-DEC-2000: 2000US-0251988.
PR 05-DEC-2000: 2000US-0256719.
PR 06-DEC-2000: 2000US-0251479.
PR 08-DEC-2000: 2000US-0251856.
PR 08-DEC-2000: 2000US-0251868.
PR 08-DEC-2000: 2000US-0251869.
PR 08-DEC-2000: 2000US-0251894.
PR 08-DEC-2000: 2000US-0251990.
PR 11-DEC-2000: 2000US-0254097.
PR 05-JAN-2001: 2001US-0254678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM.
XX
XX WPI: 2001-483426/52.
XX

```

```

PT Nucleic acids encoding human immune/haematopoietic anti-
PT useful for preventing, diagnosing and/or treating cancer
PT metastasis -
XX
XX
PS Disclosure: SEQ ID NO: 25604; 307bp; - sequence listing
XX
XX AAK45951 to AAK6702 encode the human immune/haematopoie
XX amino acid sequences given in AAK62170 to AAK61921. (1) i
XX activity, and can be used in gene therapy and vaccine pr
XX proteins and polynucleotides may be used in the preventi
XX treatment of diseases associated with inappropriate (1) i
XX example, they may be used to treat disorders associated a
XX expression by rectifying mutations or deletions in a pat
XX that affect the activity of (1) by expressing inactive pr
XX supplement the patients own production of (1). Additiona
XX polynucleotides may be used to produce the secreted (1) i
XX the nucleic acids into a host cell and culturing the cel
XX protein. (1) proteins and polynucleotides may be used to
XX diagnose and treat immune/haematopoietic-related disease
XX cancer and cancer metastases of haematopoietic-derived c
XX to AAK67634 represent human immune/haematopoietic anti-b
XX sequences from the present invention. AAK54942 to AAK549
XX represent sequences used in the exemplification of the p
XX
XX Sequence 319 bp: 96 A; 77 C; 56 G; 90 T; 0 other:
XX
XX Query Match 1,387 score 20; DB 22; Length 3
XX Best local Similarity 100.0%; prod. No. 18;
XX Matches 20; Conservative 0; Mismatches 0; Indels
XX
XX QY 1480 HGGCTGTTCTGAGTGTGA 1479
XX Db 238 HGGCTGTTCTGAGTGTGA 219
XX
XX RESULT 42
XX AAK75083/C
XX ID AAK75083 standard, DNA: 319 bp.
XX
XX AC AAK75083;
XX
XX U1 07-NOV-2001 (first entry)
XX
XX DE Human immune/haematopoietic antigen genomic sequence SPL
XX
XX KW Human, immune, haematopoietic; immune/haematopoietic ant;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX US Homo sapiens.
XX
XX IN Wc200157162 A2.
XX
XX PD 09-AUG-2001.
XX
XX XX
XX 17-JAN-2001, 2001WD-0501354.
XX
XX FE 41-JAN-2000; 2000US-0179065.
XX XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184654.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 14-JUL-2000; 2000US-0217496.
XX 26-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.

```


CC to AAK67694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention AAK5442 to AAK5460 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
XX
SO Sequence 319 BP: 96 A; 77 C; 56 G; 90 T; 0 other;

Query Match 1.3%; Score 20; DB 22; Length 319;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1480 TTGCTGTTTGACTGTA 1499
|||||
DB 248 TTGCTGTTTGACTGTA 219

RESULT 44
AA094106/c
ID AA094106 standard; DNA: 692 BP.
XX
AC AA094106;
XX
DT 23-MAY-2001 (first entry)
XX
DE Primer specific for DNA encoding secretory/membrane protein Shy ID 540.
XX
KM Human; secretory protein; membrane protein; vaccine; gene therapy;
KM rheumatoid arthritis; diabetes; PCR primer, SS.
XX
OS Synthetic.
XX
PN EP1067182-A2.
XX
PD 10-JAN-2001.
XX
PF 07-JUL-2000; 2000EP-0114090.
XX
PR 08-JUL-1999; 99JP-0194179.
PR 11-JAN-2000; 2000JP-0118775.
PR 02-MAY-2000; 2000JP-0183766.
XX
XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;
XX
XX WPI: 2001-093989/11.
XX
XX Nucleic acids encoding secretory proteins/membrane proteins, useful in
XX gene therapy or as candidate target molecules in drug development -
XX
XX
XX Claim 4; SEQ ID 540; 609pp + CD ROM; English.

CC This invention relates to nucleic acid sequences AAF93744 - AAF93916
CC which encode human secretory or membrane proteins represented by
CC AAB88317 - AAB88419 included in the invention are primers
CC AAF93917 - AAF94295 and AAF62232 - AAF62235 which are used to isolate the
CC cDNA sequences of the invention. The invention also includes methods for
CC the production of antibodies directed against the proteins, and cDNA
CC sequences, which can be used in vaccines. The polynucleotide sequences
CC can be used in gene therapy. The polynucleotide sequences and the
CC proteins they encode may be used in the prevention, treatment and
CC diagnosis of diseases associated with inappropriate secretory
CC protein/membrane protein expression. The nucleic acids and complementary
CC sequences may also be used as DNA probes in diagnostic assays
CC (e.g. polymerase chain reactions (PCR)) to detect and quantitate the
CC presence of similar nucleic acid sequences in samples. They may also be
CC used to study the expression and function of secretory proteins/membrane
CC polypeptides and their role in metabolism. The polypeptides may be used
CC as antigens in the production of antibodies against them and in assays to
CC identify modulators (agonists and antagonists) of expression and
CC activity. The antibodies and antagonists may also be used as therapeutic
CC agents to down regulate expression and activity. The antibodies may also
CC be used as diagnostic agents for detecting the presence of the
CC polypeptides in samples (e.g. by enzyme linked immunosorbent assay

CC (ELISA). Examples of diseases which may be treated include
CC arthritis and diabetes.
XX
XX
SO Sequence 692 BP: 149 A; 192 C; 192 G; 156 T; 3 other;

Query Match 1.3%; Score 20; DB 22; Length 692;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0;
QY 366 TGGTGGGCTGCTGGGCTG 485
|||||
DB 223 TGGTGGGCTGCTGGGCTG 204

RESULT 44
AA087427
ID AA087427 standard; CDNA: 1983 BP.
XX
AC AA087427;
XX
DT 16-NOV-1995 (first entry)
XX
DE Rat GRK cDNA clone B24.
XX
KM Polymerase chain reaction; PCR; amplification; primer; isolated
KM G-protein-coupled receptor kinase; GRK6; catalytic domain
KM Rhodopsin kinase; RK; beta adrenergic receptor kinase; B
KM conserved amino acid; subdomain II; subdomain VII; graft
KM signal transduction; immune system; inflammation; human;
XX
OS Rattus rattus.
XX
FH Key
FH CDS Location/Qualifiers
FT 13..174
FT /*tag= a
FT /product= GRK6
XX
PN W04507440 A.
XX
PD 23-MAR-1995.
XX
PF 16-SEP-1994; 94WO-0510487.
XX
PR 17-SEP-1993; 93US-0124942.
PR 31-MAR-1994; 94US-0221817.
XX
XX (ICOS-) ICOS CGRP.
XX
PI Chantry D, Gray PW, Hoekstra MF;
XX
XX WPI: 1995-131358/17.
XX P-PSDB: AAR71175.

XX New G protein-coupled receptor kinase, GRK6, and correspond-
XX ing cDNA sequences are disclosed. The cDNA sequences are used
XX in the production of GRK6 and for the immunisation, treat-
XX ment and diagnosis of diseases of the immune system and inflammatory
XX diseases.
XX
XX Example 3; Page 48-50; 71pp; English.

CC This sequence represents the cDNA clone, B24, which encodes
CC rat G protein-coupled receptor kinase, GRK6. This fragment
CC from a rat thymus library in lambda ZAP, using the PCR to
CC amplified from rat T cell cDNA using the primers given
CC The rat and human GRK6 sequences exhibit 95% identity at
CC level. Peptide fragments of GRK6 may be used in immunisation
CC and quantification of GRK6. They may also be used for
CC activities of GRK6, esp. those involved in signal trans-
CC Modulators of GRK6 may be useful in treatment of disease
CC physiological conditions of the immune system, such as
XX
XX Sequence 1983 BP: 453 A; 528 C; 594 G; 403 T; 0 other;

Query Match 1.3%; Score 20; DB 16; Length

Best local similarity: 100.0%, Prod. No.: 17,
Matches: 20, Conservative: 0, Mismatches: 0, Indels: 0, Gaps: 0,

QY 562 CTTAGCATGTGACCGGCTGA 611
|||||
DB 438 CCAAGAGCTGACCGGCTGA 457

RESULT 45
AAF93916/C
ID AAF93916 standard; cDNA; 2646 BP.

AC AAF93916;

DI 23-MAY-2001 (first entry)

DE Human cDNA encoding a membrane or secretory protein clone pSEC0265.

KA Human; secretory protein; membrane protein; vaccinia; gene therapy;

KA Rheumatoid arthritis; diabetes; ss.

OS Homo sapiens.

PN EP1067182-A2.

PE 10-JAN-2001.

PF 07-MAY-2000; 2000EP-0114090.

PR 08-JUL-1999; 99JP-0194179.

PR 11-JAN-2000; 2000JP-0118775.

PR 02-MAY-2000; 2000JP-0183766.

XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Kawai Y, Sugiyama T, Hayashi K;

DR WPI: 2001-093989/11.

PS P-PSDB: AAB88489.

PT Nucleic acids encoding secretory proteins/membrane proteins, useful in

XX gene therapy or as candidate target molecules in drug development.

XX Claim 1: SEQ ID 349; 609pp + CD ROM; English.

XX This invention relates to nucleic acid sequences AAF93744 - AAF93916

XX which encode human secretory or membrane proteins represented by

XX AAB88317 - AAB88419. Included in the invention are primers

XX AAF93917 - AAF94295 and AAF6232 - AAF6235 which are used to isolate the

XX cDNA sequences of the invention. The invention also includes methods for

XX the production of antibodies directed against the proteins, and cDNA

XX sequences, which can be used in vaccines. The polynucleotide sequences

XX can be used in gene therapy. The polynucleotide sequences and the

XX proteins they encode may be used in the prevention, treatment and

XX diagnosis of diseases associated with inappropriate secretory

XX protein/membrane protein expression. The nucleic acids and complementary

XX sequences may also be used as DNA probes in diagnostic assays

XX (e.g. polymerase chain reactions (PCR)) to detect and quantitate the

XX presence of similar nucleic acid sequences in samples. They may also be

XX used to study the expression and function of secretory protein/membrane

XX polypeptides and their role in metabolism. The polypeptides may be used

XX as antigens in the production of antibodies against them and in assays to

XX identify modulators (agonists and antagonists) of expression and

XX activity. The antibodies and antagonists may also be used as therapeutic

XX agents to down regulate expression and activity. The antibodies may also

XX be used as diagnostic agents for detecting the presence of the

XX polypeptides in samples (e.g. by enzyme linked immunosorbent assay

XX (ELISA)). Examples of diseases which may be treated include rheumatoid

XX arthritis and diabetes.

XX SEQIDNO: 2646 PP: 698 A: 615 C: 612 G: 721 T: 0 other;

XX Query Match: 1.39; Score 20; DB 22; Length 2646;

Best local similarity: 100.0%, Prod. No.: 17,
Matches: 20, Conservative: 0, Mismatches: 0,

QY 366 TCGTCTGAGCTGCTGGTGGG 385
|||||
DB 223 TCGTCTGAGCTGCTGGTGGG 204

Search completed: December 7, 2002, 09:38:21
Job time : 348 secs


```

|||||
Db 313 GATTGGCCAGAAATCCAGAGCTGACGAGATCTACAGAGCTGACCTGGATGGAGCTT 372
QY 617 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 676
Db 373 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
QY 677 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 736
Db 433 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
QY 737 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
Db 493 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 530

```

```

RESULT 2
US-09-591-435-9
: Sequence 9, Application US/09591435
: Patent No. 6280953
: GENERAL INFORMATION:
: APPLICANT: MESSIER, WALTER
: APPLICANT: SIKELA, JAMES M
: TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
: TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
: FILE REFERENCE: GPN.200.2
: CURRENT APPLICATION NUMBER: US/09/591,435
: CURRENT FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/591,435
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/240,915
: PRIOR FILING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: 60/073,263
: PRIOR FILING DATE: 1998-01-30
: PRIOR APPLICATION NUMBER: 60/098,987
: PRIOR FILING DATE: 1998-09-02
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9
: LENGTH: 1212
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-591-435-9

```

```

Query Match 7.2%; Score 111; DB 4; Length 1212;
Best Local Similarity 100.0%; Pred. No. 3,7e-46;
Matches 111, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

```

```

QY 449 ATCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
Db 229 ATCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 288
QY 509 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 559
Db 289 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 339

```

```

RESULT 3
US-09-591-435-10
: Sequence 10, Application US/09591435
: Patent No. 6280953
: GENERAL INFORMATION:
: APPLICANT: MESSIER, WALTER
: APPLICANT: SIKELA, JAMES M
: TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
: TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
: FILE REFERENCE: GPN.200.2
: CURRENT APPLICATION NUMBER: US/09/591,435
: CURRENT FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/591,435
: PRIOR FILING DATE: 2000-06-09

```

```

: PRIOR APPLICATION NUMBER: 09/240,915
: PRIOR FILING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: 60/073,263
: PRIOR FILING DATE: 1998-01-30
: PRIOR APPLICATION NUMBER: 60/098,987
: PRIOR FILING DATE: 1998-09-02
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 1212
: TYPE: DNA
: ORGANISM: Pan troglodytes
US-09-591-435-10

```

```

Query Match 7.2%; Score 111; DB 4; Length 1212;
Best Local Similarity 100.0%; Pred. No. 3,7e-46;
Matches 111, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

```

```

QY 449 ATCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
Db 229 ATCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 288
QY 509 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 559
Db 289 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 339

```

```

RESULT 4
US-09-517-605-1
: Sequence 1, Application US/09517605
: Patent No. 6391567
: GENERAL INFORMATION:
: APPLICANT: Littman, Dan R.
: APPLICANT: Kwon, Douglas S.
: APPLICANT: van Kooijk, Yvette
: APPLICANT: Geljtenbeck, Theo
: TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF REPRODUCTION ENTRY INTO
: TITLE OF INVENTION: CELLS
: FILE REFERENCE: 1049-1-017
: CURRENT APPLICATION NUMBER: US/09/517,605
: CURRENT FILING DATE: 2000-03-02
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1312
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (42)..(1453)
US-09-517-605-1

```

```

Query Match 7.2%; Score 111; DB 4; Length 1312;
Best Local Similarity 100.0%; Pred. No. 3,7e-46;
Matches 111, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

```

```

QY 449 ATCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 508
Db 270 ATCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 329
QY 509 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 559
Db 330 CTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 380

```

```

RESULT 5
US-09-591-435-11
: Sequence 11, Application US/09591435
: Patent No. 6280953
: GENERAL INFORMATION:
: APPLICANT: MESSIER, WALTER
: APPLICANT: SIKELA, JAMES M
: TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE

```

1 TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
2 TITLE OF INVENTION: AND MEDICAL CONDITIONS
3 FILE REFERENCE: GENO. 200.2
4 CURRENT APPLICATION NUMBER: US/09/591,435
5 CURRENT FILING DATE: 2000-06-09
6 PRIOR APPLICATION NUMBER: 09/591,435
7 PRIOR FILING DATE: 2000-06-09
8 PRIOR APPLICATION NUMBER: 09/240,915
9 PRIOR FILING DATE: 1999-01-29
10 PRIOR APPLICATION NUMBER: 60/074,263
11 PRIOR FILING DATE: 1998-01-30
12 PRIOR APPLICATION NUMBER: 60/098,987
13 PRIOR FILING DATE: 1998-09-02
14 NUMBER OF SFO ID NOS: 13
15 SOFTWARE: Patent In Ver. 2.0
16 SEQ ID NO: 11
17 LENGTH: 1212
18 TYPE: DNA
19 ORGANISM: Gorilla gorilla
20 US-09-591-435-11

Query Match 4.18; Score 63; DB 4; Length 1212;
Best Local Similarity 100.0%; Pred. No. 56-22;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 860 UGACATCTTCGACGAGACTGTACTGCTACTGCTACTGCTGAGGAGGAGCTGGACGAC 919
DB 778 TGACATCTTCGACGAGACTGTACTGCTACTGCTACTGCTGAGGAGGAGCTGGACGAC 837

UY 929 TCC 922
DB 838 TCC 840

RESULT 6
US-08-221-817-21
Sequence 21, Application US/08221817
Patent No. 5532151
GENERAL INFORMATION:
APPLICANT: Chantry, David
APPLICANT: Gray, Patrick W.
APPLICANT: Hockstra, Merle F.
TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/221,817
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
ATTORNEY/AGENT INFORMATION:
NAME: No. 5532151and, Grete F.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856

1 INFORMATION FOR SEQ ID NO: 21:
2 SEQUENCE CHARACTERISTICS:
3 LENGTH: 1983 base pairs
4 TYPE: nucleic acid
5 STRANDEDNESS: single
6 TOPOLOGY: linear
7 MOLECULE TYPE: cDNA
8 FEATURE:
9 NAME/KEY: CDS
10 LOCATION: 13..1740
11 US-08-221-817-21

Query Match 1.38; Score 20; DB 1; Length 1983;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0;

UY 552 CGAGAGCTGACCGGCTGA 911
DB 438 CTAGAGCTGACCGGCTGA 457

RESULT 7
US-08-454-439-21
Sequence 21, Application US/08454439
Patent No. 5591618

GENERAL INFORMATION:
APPLICANT: Chantry, David
APPLICANT: Gray, Patrick W.
APPLICANT: Hockstra, Merle F.
TITLE OF INVENTION: A No. 5591618el G Protein-Coupled R
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,439
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/221,817
FILING DATE: 31-MAR-1994
APPLICATION NUMBER: 08/123,932
FILING DATE: 17 SEP 1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5591618and, Grete F.
REGISTRATION NUMBER: 45,402
REFERENCE/DOCKET NUMBER: 31981
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 1983 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1740

US-08-454-449-21

Query Match: 1.3%; Score 20; DB 1; Length 1983;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

U7 592 CAGAGAGTACCGGCTGA 611
|||||
DB 438 CAGAGAGTACCGGCTGA 457

RESULT 8

PCT-US94-10487-21

Sequence 21. Application PC/TUS9410487

GENERAL INFORMATION:

APPLICANT: ICOS Corporation

TITLE OF INVENTION: A Novel G Protein Coupled Receptor

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &

STREET: 6300 Sears Tower, 213 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/10487

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/221,817

FILING DATE: 31 MAR 1994

CLASSIFICATION:

APPLICATION DATA:

APPLICATION NUMBER: 08/223,932

FILING DATE: 17 SEP 1993

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Noland, Greca E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 27866/1981

TELEPHONE: (312) 474-6400

TELEFAX: (312) 474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 1983 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 13..1740

PCT-US94-10487-21

Query Match: 1.3%; Score 20; DB 5; Length 1983;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

U7 592 CAGAGAGTACCGGCTGA 611
|||||
DB 438 CAGAGAGTACCGGCTGA 457

RESULT 9

US-09-222-575-2976

Sequence 20. Application US/09222575

Patent No. 636547

GENERAL INFORMATION:

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

APPLICANT: Yung, Jiang

Post Local Similarity 100.0%; Pred. No. 20;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1525 TGTACTGCAAAAAA 1542
|||||
DB 155 TGTACTGCAAAAAA 172

RESULT 11

US-08-938-291A-3

Sequence 3, Application US/08938291A

Patent No. 6117673

GENERAL INFORMATION:

APPLICANT: Lev, Sima

APPLICANT: Plowman, Gregory D.

APPLICANT: Schlessinger, Joseph

TITLE OF INVENTION: RGDH PROTEINS AND RELATED

TITLE OF INVENTION: PRODUCTS AND METHODS

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: 633 West Fifth Street

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: U.S.A.

ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 MB

COMPUTER: IBM Compatible

OPERATING SYSTEM: IBM P.C. DOS 5.0

SOFTWARE: FastSeq

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/938,291A

FILING DATE: September 26, 1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/027,337

FILING DATE: October 11, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Walburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 228/172

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5020 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: Linear

US-08-938-291A-3

Query Match 1 2% Score 18, NR 3, Length 5020;

Best Local Similarity 100.0%; Pred No. 21;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1429 TACAACTCTTCACGCC 1446
|||||
DB 2344 TACAACTCTTCACGCC 2361

RESULT 12

US-09-064-411A-33/C

Sequence 33, Application US/09064411A

Patent No. 6331664

GENERAL INFORMATION:

APPLICANT: Rubin-Wilson, Beth

APPLICANT: Guo, Lining

APPLICANT: Skokut, Tom

APPLICANT: Young, Scott

APPLICANT: Folkerts, Otto

APPLICANT: Armstrong, Katherine

APPLICANT: Cowen, Neil M.

TITLE OF INVENTION: Nucleotide Sequences of Major Oligo-

TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thio-

TITLE OF INVENTION: Their Use in the Regulation of Fatty

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dow Agrosciences Patent Department

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: Indiana

COUNTRY: USA

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25" disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Process #10, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/064,411A

FILING DATE: 22-APR-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/045,827

FILING DATE: 05-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Horucki, Andrea L.

REGISTRATION NUMBER: 34651

REFERENCE/DOCKET NUMBER: 50444

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-337-4846

TELEFAX: 317-337-4847

INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:

LENGTH: 324 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..324

US-09-064-411A-33

Query Match 1.1% Score 17, OR 4, Length 324;

Best Local Similarity 100.0%; Pred. No. 63;

Matches 17; Conservative 0; Mismatches 0; Indels 4; Gaps 0;

QY 1114 CAGTGGAGACCAACA 1130
|||||
DB 100 CAGTGGAGACCAACA 84

RESULT 13

US-08-905-223-79

Sequence 79, Application US/08905223

Patent No. 6228929

GENERAL INFORMATION:

APPLICANT: Edwards, Jean-Baptiste D.

APPLICANT: Duelet, Aymeric

APPLICANT: Lacroix, Bruno

TITLE OF INVENTION: 5' PSIS FOR SELECTED PROTEINS

NUMBER OF SEQUENCES: 503

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Matveus, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101-3505

COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Win95
SOFTWARE: Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/905,723
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Israelson, Ned A.
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo Sapiens
TISSUE TYPE: Brain
FEATURE:
NAME/KEY: siq.peptide
LOCATION: 61..372
IDENTIFICATION METHOD: Von Heijne matrix
OTHER INFORMATION: score 6.6
OTHER INFORMATION: seq HPSHVVWHPHWA/00
US-09-905-223-79

Query Match: 1.1% Score 17; DB 4; Length 402;
Best Local Similarity 100.0%; Pred No 64;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 264 AAACATGACTGACTGC 280
UU 56 AAACATGACTGACTGC 72

RESULT 14
US-09-064-411A-43/c
Sequence 43, Application US/09064411A
Patent No. 6331664
GENERAL INFORMATION:
APPLICANT: Rubin-Wilson, Beth
APPLICANT: Guo, Lining
APPLICANT: Skokut, Tom
APPLICANT: Young, Scott
APPLICANT: Folkerts, Otto
APPLICANT: Armstrong, Katherine
APPLICANT: Cowen, Neil M.
TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP
TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and
NUMBER OF INVENTION: Their Use In The Regulation of Fatty Acid Content of Oil
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,411A
FILING DATE: 22-APR-1998

```

```

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 6,0/045,827
FILING DATE: 05-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Potucki, Andrea I.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50433
TELECOMMUNICATION INFORMATION:
TELEPHONE: 417-337-4846
TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 987 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 11..970
US-09-064-411A-43

Query Match: 1.1% Score 17; DB 4; Length 987;
Best Local Similarity 100.0%; Pred No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 1114 CAGTGGAGAACCCACAA 1130
UU 257 CAGTGGAGAACCCACAA 241

RESULT 15
US-09-064-411A-36/c
Sequence 36, Application US/09064411A
Patent No. 6331664
GENERAL INFORMATION:
APPLICANT: Rubin-Wilson, Beth
APPLICANT: Guo, Lining
APPLICANT: Skokut, Tom
APPLICANT: Young, Scott
APPLICANT: Folkerts, Otto
APPLICANT: Armstrong, Katherine
APPLICANT: Cowen, Neil M.
TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP
TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and
NUMBER OF INVENTION: Their Use In The Regulation of Fatty Acid Content of Oil
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,411A
FILING DATE: 22-APR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 6,0/045,827
FILING DATE: 05-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Potucki, Andrea I.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50433
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-337-4846

```

TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 996 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..996
US-09-064-411A-38
Query Match 1.1%; Score 17; DB 4; Length 996;
Host Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1114 CAGTGGAGAACCCACACA 1130
|||||
DB 283 CAGTGGAGAACCCACACA 267

RESULT 16
US-09-064-411A-36/C
Sequence 36, Application US/09064411A
Patent No. 631664
GENERAL INFORMATION:
APPLICANT: Rubin-Wilson, Beth
APPLICANT: Guo, Lining
APPLICANT: Skokut, Tom
APPLICANT: Young, Scott
APPLICANT: Folkerts, Otto
APPLICANT: Armstrong, Katherine
TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP
TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and
TITLE OF INVENTION: Their Use In The Regulation of Fatty Acid Content of Oil
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,411A
FILING DATE: 22-APR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/045,827
FILING DATE: 05-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Borucki, Andrea T.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50433
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-337-4846
TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1098 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS

LOCATION: 1..1098
FEATURE:
NAME/KEY: mal_peptide
LOCATION: 103..1098
US-09-064-411A-36
Query Match 1.1%; Score 17; DB 4; Length 1098;
Host Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 1114 CAGTGGAGAACCCACACA 1130
|||||
DB 385 CAGTGGAGAACCCACACA 369

RESULT 17
US-09-064-411A-47
Sequence 47, Application US/09064411A
Patent No. 631664
GENERAL INFORMATION:
APPLICANT: Rubin-Wilson, Beth
APPLICANT: Guo, Lining
APPLICANT: Skokut, Tom
APPLICANT: Young, Scott
APPLICANT: Folkerts, Otto
APPLICANT: Armstrong, Katherine
TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP
TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and
TITLE OF INVENTION: Their Use In The Regulation of Fatty Acid Content of Oil
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dow Agrosciences Patent Department
STREET: 9330 Zionsville Road
CITY: Indianapolis
STATE: Indiana
COUNTRY: USA
ZIP: 46268
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/064,411A
FILING DATE: 22-APR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/045,827
FILING DATE: 05-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Borucki, Andrea T.
REGISTRATION NUMBER: 33651
REFERENCE/DOCKET NUMBER: 50433
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-337-4846
TELEFAX: 317-337-4847
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: YES
US-09-064-411A-47
Query Match 1.1%; Score 17; DB 4; Length 1140;
Host Local Similarity 100.0%; Pred. No. 65;
Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 1114 CAGTGGAGAACCCACACA 1130
|||||

1b 745 CAGTGGAGAGCCGACCA 751

RESULT 18

US-09-538-871-3

Sequence 3, Application US/09518871

Patent No. 6358718

GENERAL INFORMATION:

APPLICANT: Olson, Mark

APPLICANT: Szedeni, Attila

TITLE OF INVENTION: METHODS FOR STABILIZATION AND

FILE OF INVENTION: RENATURATION OF PROTEINS USING NUCLEOLAR PROTEINS P24

CURRENT APPLICATION NUMBER: US/09/538,871

EARLIER FILING DATE: 2000-03-30

EARLIER APPLICATION NUMBER: 60/126,910

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO: 3

LENGTH: 1164

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE: OTHER INFORMATION: Description of artificial sequence: No 6358718c

US-09-538-871-3

Query Match 1.1% Score 17; DB 4; Length 1164;

Best Local Similarity 100.0%; Prev. No. 65;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1454 CTTGAGGCTGTGAAGCA 1370

1b 992 CTTGAGGCTGTGAAGCA 1098

RESULT 19

US-09-064-411A-35/C

Sequence 35, Application US/09064411A

Patent No. 6331664

GENERAL INFORMATION:

APPLICANT: Rudin-Wilson, Beth

APPLICANT: Guo, Lining

APPLICANT: Skokul, Tom

APPLICANT: Young, Scott

APPLICANT: Folkerst, Otto

APPLICANT: Armstrong, Katherine

TITLE OF INVENTION: Nucleotide Sequences of Maize Oleoyl-ACP

TITLE OF INVENTION: Thioesterase and Palmitoyl-ACP Thioesterase Genes and

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dow AgroSciences Patent Department

STREET: 9330 Zionsville Road

CITY: Indianapolis

STATE: Indiana

COUNTRY: USA

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/064,411A

FILING DATE: 22-APR-1998

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/045,827

FILING DATE: 05-MAY-1997

ATTORNEY/AGENT INFORMATION:

NAME: Horucki, Andrea T.

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 50433

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-347-4846

TELEFAX: 317-347-4847

INFORMATION FOR SEQ ID NO: 45;

SEQUENCE CHARACTERISTICS:

LENGTH: 1287 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

FEATURE:

NAME/KEY: CUS

LOCATION: 38..1135

FEATURE:

NAME/KEY: mat_peptide

LOCATION: 140..1135

US-09-064-411A-35

Query Match 1.1% Score 17; DB 4; Length 1287;

Best Local Similarity 100.0%; Prev. No. 65;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1114 CAGTGGAGAGCCGACCA 1130

Db 422 CAGTGGAGAGCCGACCA 406

RESULT 20

US-08-581-148C-15/C

Sequence 15, Application US/08581148C

Patent No. 6060644

GENERAL INFORMATION:

APPLICANT: Schmale, Patrick S.

APPLICANT: Robertson, Donald S.

APPLICANT: Hansen, Joel D.

APPLICANT: Nikolau, Basil J.

APPLICANT: Xu, Xiaojie

TITLE OF INVENTION: ISOLATION AND USE OF COTICULIN LIPID

TITLE OF INVENTION: GENES

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: Levitz, Velt & Mayer, Ltd.

STREET: Two Prudential Plaza, Suite 4900

CITY: Chicago

STATE: IL

COUNTRY: US

ZIP: 60601-6/80

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/581,148C

FILING DATE: 29-DEC-1995

CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:

NAME: Latchford, Carol

REGISTRATION NUMBER: 35243

REFERENCE/DOCKET NUMBER: 71380

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 616-5600

TELEFAX: (312) 616-5700

INFORMATION FOR SEQ ID NO: 15;

SEQUENCE CHARACTERISTICS:

LENGTH: 1903 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLBIOLE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1665
US-08-581-148C-15

Query Match 1.18; Score 17; DB 3; Length 1903;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 43 GGAGCTCTTCCTCTGT 49
|||||
DB 567 GGAGCTCTTCCTCTGT 551

RESULT 21
US-09-305-640-1
Sequence 1, Application US/09305640B
Patent No. 6255468
GENERAL INFORMATION:
APPLICANT: SmithKline Beecham plc
TITLE OF INVENTION: No 6255468a1 Compounds
FILE REFERENCE: GP30124
CURRENT APPLICATION NUMBER: US/09/305,640B
CURRENT FILING DATE: 1999-05-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2560
TYPE: DNA
ORGANISM: Homo sapiens
US-09-305-640-1

Query Match 1.18; Score 17; DB 4; Length 2560;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 844 CCAGCTGCCCAAGGACT 860
|||||
DB 2499 CCAGCTGCCCAAGGACT 2515

RESULT 22
US-09-141-206-1
Sequence 1, Application US/09141206
Patent No. 6187559
GENERAL INFORMATION:
APPLICANT: Steed, Paul M
TITLE OF INVENTION: Amino Acid Sequence of Human PLD2A
Patent No. 6187559
FILE REFERENCE: 4-30148/P1/0701954/P
CURRENT APPLICATION NUMBER: US/09/141,206
CURRENT FILING DATE: 1998-08-27
EARLIER APPLICATION NUMBER: 60/057,802
EARLIER FILING DATE: 1997-08-28
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3388
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: CDS
LOCATION: (162)..(2963)
OTHER INFORMATION: Human PLD2
US-09-141-206-1

Query Match 1.18; Score 17; DB 4; Length 3388;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Caps 0;
QY 981 AACTCTACAGCTGCA 999

|||||
DB 1107 AACTCTACAGCTGCA 1124

RESULT 23
US-09-600-776-9/c
Sequence 9, Application US/09600776
Patent No. 6326168
GENERAL INFORMATION:
APPLICANT: Yamaguchi Pharmaceutical Co., Ltd.
TITLE OF INVENTION: A novel potassium channel protein
FILE REFERENCE: Y9903-PCF
CURRENT APPLICATION NUMBER: US/09/600,776
CURRENT FILING DATE: 2000-07-21
PRIOR APPLICATION NUMBER: JP P1598-011434
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: JP P1598-446198
PRIOR FILING DATE: 1998-12-04
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9
LENGTH: 3715
TYPE: DNA
ORGANISM: Rattus sp.
US-09-600-776-9

Query Match 1.18; Score 17; DB 4; Length 3715;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 919 CTCCTACACCGCTGCG 935
|||||
DB 2825 CTCCTACACCGCTGCG 2809

RESULT 24
US-08-769-309A-4/c
Sequence 4, Application US/08769309A
Patent No. 5741890
GENERAL INFORMATION:
APPLICANT: Scott, John J.
APPLICANT: Nauert, Brian J.
APPLICANT: Klauack, Theresa M.
TITLE OF INVENTION: Protein Binding Domains of Grp78
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Perslein, Murray & Box
STREET: 6300 Sears Tower/233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/769,309A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5741890and, Giga E.
REGISTRATION NUMBER: 45,402
REFERENCE/DOCKET NUMBER: 27866/34451
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6605 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 192..5531
US-08-769-409A-4

Query Match 1.1%; Score 17; DB 1; Length 6605;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 GGTGGAGTCTCTGCG 383
DB 846 GGTGGAGTCTCTGCG 830

RESULT 25

US-08-994-570-4/c
Sequence 4, Application US/08994570
Patent No. 6090929

GENERAL INFORMATION:

APPLICANT: Scott, John D.

APPLICANT: Nauer, Brian J.

APPLICANT: Klauk, Theresa M.

TITLE OF INVENTION: Protein Binding Domains of Gravin

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borum

STREET: 6300 Sears Tower/233 South Wacker Drive

CITY: Chicago

STATE: Illinois

COUNTRY: United States of America

ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/994,570

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: No. 6090929and, Greta E.

REGISTRATION NUMBER: 35,302

REFERENCE/DOCKET NUMBER: 27866/4441

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-474-6300

TELEFAX: 312-474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 6605 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 192..5531

US-08-994-570-4

Query Match 1.1%; Score 17; DB 3; Length 6605;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 GGTGGAGTCTCTGCG 383
DB 846 GGTGGAGTCTCTGCG 830

RESULT 26

US-08-665-259-19/c

Sequence 19, Application US/08665259

Patent No. 6028173

GENERAL INFORMATION:

APPLICANT: Landes, Gregory M.

APPLICANT: Burn, Timothy C.

APPLICANT: Connors, Timothy D.

APPLICANT: Dackowski, William R.

APPLICANT: Van Raay, Terence J.

APPLICANT: Klinder, Katherine W.

TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,

NUMBER OF SEQUENCES: 73

CORRESPONDENCE ADDRESS:

ADDRESS: GENZYME CORPORATION

STREET: One Mountain Road

CITY: Framingham

STATE: Massachusetts

COUNTRY: United States of America

ZIP: 01701

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/665,259

FILING DATE: 17-JUN-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Dugan, Deborah A.

REGISTRATION NUMBER: 37,415

REFERENCE/DOCKET NUMBER: 165-9.1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (508) 872-8400

TELEFAX: (508) 872-5415

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 6803 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-665-259-19

Query Match 1.1%; Score 17; DB 3; Length 6803;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 GGTGGAGTCTCTGCG 414
DB 4097 GGTGGAGTCTCTGCG 4081

US-08-762-500-19/c

Sequence 19, Application US/08762500

Patent No. 6030806

GENERAL INFORMATION:

APPLICANT: Landes, Gregory M.

APPLICANT: Burn, Timothy C.

APPLICANT: Connors, Timothy D.

APPLICANT: Dackowski, William R.

APPLICANT: Van Raay, Terence J.

APPLICANT: Klinder, Katherine W.

TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,

NUMBER OF SEQUENCES: 83

CORRESPONDENCE ADDRESS:

ADDRESS: GENZYME CORPORATION

STREET: One Mountain Road

CITY: Framingham

STATE: Massachusetts

```

: COUNTRY: United States of America
: ZIP: 01701
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/762,500
: FILING DATE: 09-DEC-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/665,276
: FILING DATE: 17-JUN-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCI/0596/10469
: FILING DATE: 17-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Dugan, Deborah A.
: REGISTRATION NUMBER: 47,315
: REFERENCE/DOCKET NUMBER: IG5-9.3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (508) 872-8400
: TELEFAX: (508) 872-5415
: INFORMATION FOR SEQ. ID NO.: 19
: SEQUENCE CHARACTERISTICS:
: LENGTH: 6803 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-762-500-19

Query Match 1.18; Score 17, DB 3, Length 6803,
Best Local Similarity 100.0%; Pred. No. 69;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 398 GTGTCCAGTCCCGCAG 414
DB 4097 GTGTCCAGTCCCGCAG 4081

RESULT 28
: US-08-222-617A-26
: Sequence 26, Application US/08222617A
: Patent No. 5882879
: GENERAL INFORMATION:
: APPLICANT: Veenstra, Annemarie E.
: APPLICANT: Martin, Juan F.
: APPLICANT: Garcia, Bruno D.
: APPLICANT: Gutierrez, Santiago
: APPLICANT: Barredo, Jose L.
: APPLICANT: Von Boehren, Hans
: APPLICANT: Palissa, Harriet
: APPLICANT: Montenegro, Eduardo P.
: TITLE OF INVENTION: A Method for Influencing Beta-lactam
: TITLE OF INVENTION: Antibiotic Production and for Isolation of Large
: TITLE OF INVENTION: Quantities of ACV Synthetase
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
: STREET: 300 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/222,617A
: FILING DATE: 04-APR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: REFERENCE/DOCKET NUMBER: 97,157
: INFORMATION FOR SEQ. ID NO.: 1
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12364 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Penicillium chrysogenum
: FEATURE:
: NAME/KEY: GDS
: LOCATION: 264..11600

```

```

: APPLICATION NUMBER: US/08/222,617A
: FILING DATE: 04-APR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: REFERENCE/DOCKET NUMBER: 97,157
: INFORMATION FOR SEQ. ID NO.: 26;
: SEQUENCE CHARACTERISTICS:
: LENGTH: 11444 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: US-08-222-617A-26

Query Match 1.18; Score 17, DB 2, Length 114
Best Local Similarity 100.0%; Pred. No. 70;
Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 668 ATTCTTAAAGAGAA 984
DB 7740 ACTGCTGAGAGACGAA 7756

RESULT 29
: US-08-222-617A-1
: Sequence 1, Application US/08222617A
: Patent No. 5882879
: GENERAL INFORMATION:
: APPLICANT: Veenstra, Annemarie E.
: APPLICANT: Martin, Juan F.
: APPLICANT: Garcia, Bruno D.
: APPLICANT: Gutierrez, Santiago
: APPLICANT: Barredo, Jose L.
: APPLICANT: Von Boehren, Hans
: APPLICANT: Palissa, Harriet
: APPLICANT: Montenegro, Eduardo P.
: TITLE OF INVENTION: A Method for Influencing Beta-lactam
: TITLE OF INVENTION: Antibiotic Production and for Isolation
: TITLE OF INVENTION: Quantities of ACV Synthetase
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
: STREET: 300 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: USA
: ZIP: 60606
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/222,617A
: FILING DATE: 04-APR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: REFERENCE/DOCKET NUMBER: 97,157
: INFORMATION FOR SEQ. ID NO.: 1
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12364 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Penicillium chrysogenum
: FEATURE:
: NAME/KEY: GDS
: LOCATION: 264..11600

```

```
OTHER INFORMATION: /function="enzyme"
OTHER INFORMATION: /product="AcV Synthetase"
US-08-222-617A-1
Query Match
Best Local Similarity 100.0%; Pred. No. 70;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 968 ACTGCTGAGGACGAGAA 984
|||||
Db 7740 ACTGCTGAGGACGAGAA 7756

RESULT 30
US-09-754-250 3
Sequence 3: Application US/09754250
Patent No. 6376225
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
FILE REFERENCE: C1001063
CURRENT APPLICATION NUMBER: US/09/754,250
CURRENT FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 111282
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(111282)
OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match
Host Local Similarity 1.1%; Score 17; DB 4; Length 111282;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 25 TGAATGAGGAGTCCTT 41
|||||
Db 3656 TGAATGAGGAGTCCTT 36582

RESULT 31
US-09-676-610H-24/c
Sequence 24: Application US/09676610H
Patent No. 644465
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Jacqueline Wyatt
APPLICANT: Susan M. Preler
TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HEP-1 EXPRESSION
FILE REFERENCE: R15-0138
CURRENT APPLICATION NUMBER: US/09/676,610H
CURRENT FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 182
SEQ ID NO 24
LENGTH: 169998
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon
LOCATION: (1208)...(1472)
NAME/KEY: intron
LOCATION: (1473)...(124390)
NAME/KEY: exon
LOCATION: (124391)...(124544)
NAME/KEY: intron
LOCATION: (124545)...(125409)
NAME/KEY: exon
```

```
LOCATION: (125410)...(125595)
NAME/KEY: intron
LOCATION: (125596)...(128711)
NAME/KEY: exon
LOCATION: (128712)...(128848)
NAME/KEY: intron
LOCATION: (128849)...(133400)
NAME/KEY: exon
LOCATION: (133401)...(133469)
NAME/KEY: intron
LOCATION: (133470)...(134652)
NAME/KEY: exon
LOCATION: (134653)...(134773)
NAME/KEY: intron
LOCATION: (134774)...(136116)
NAME/KEY: exon
LOCATION: (136117)...(136261)
NAME/KEY: intron
LOCATION: (136262)...(137936)
NAME/KEY: exon
LOCATION: (137937)...(138053)
NAME/KEY: intron
LOCATION: (138054)...(138637)
NAME/KEY: exon
LOCATION: (138638)...(138766)
NAME/KEY: intron
LOCATION: (138767)...(138864)
NAME/KEY: exon
LOCATION: (138865)...(138940)
NAME/KEY: intron
LOCATION: (138941)...(139765)
NAME/KEY: exon
LOCATION: (139766)...(139860)
NAME/KEY: intron
LOCATION: (139861)...(142245)
NAME/KEY: exon
LOCATION: (142246)...(142445)
NAME/KEY: intron
LOCATION: (142446)...(143605)
NAME/KEY: exon
LOCATION: (143606)...(143738)
NAME/KEY: intron
LOCATION: (143739)...(145838)
NAME/KEY: exon
LOCATION: (145839)...(145931)
NAME/KEY: intron
LOCATION: (145932)...(147385)
NAME/KEY: exon
LOCATION: (147386)...(147544)
NAME/KEY: intron
LOCATION: (147545)...(153274)
NAME/KEY: exon
LOCATION: (153275)...(153321)
NAME/KEY: intron
LOCATION: (153322)...(155088)
NAME/KEY: exon
LOCATION: (155089)...(155231)
NAME/KEY: intron
LOCATION: (155232)...(156025)
NAME/KEY: exon
LOCATION: (156026)...(156151)
NAME/KEY: intron
LOCATION: (156152)...(156826)
NAME/KEY: exon
LOCATION: (156827)...(156928)
NAME/KEY: intron
LOCATION: (156929)...(163399)
NAME/KEY: exon
LOCATION: (163400)...(163586)
LOCATION: (163587)...(163600)
US-09-676-610H-24

Query Match
Host Local Similarity 1.1%; Score 17; DB 4; Length 169998;
Best Local Similarity 100.0%; Pred. No. 74;
```



```

ADDRESS: LAHIVE & COCKFIELD LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,507
FILING DATE: 15-May-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/969,875
FILING DATE: 10-October-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 1M1-049DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..368
US-08-441-507-22

Query Match          1.0%; Score 16; DB 4; Length 368;
Best Local Similarity 100.0%; Pred. No. 26+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 319 TCTTGCCATGCGCG 314
      |||
TX 333 TCTTGCCATGCGCG 318

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Vanshoni, Darlene A.
REGISTRATION NUMBER: 36,279
REFERENCE/DOCKET NUMBER: 041.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6010
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..368
US-07-969-875A-22

Query Match          1.0%; Score 16; DB 4; Length 368;
Best Local Similarity 100.0%; Pred. No. 26+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 319 TCTTGCCATGCGCG 314
      |||
DB 333 TCTTGCCATGCGCG 318

RESULT 37
US-08-441-507-17/c
Sequence 17, Application US/08441507
Patent No. 6214358
GENERAL INFORMATION:
APPLICANT: Singh, Mohan Biji
APPLICANT: Smith, Penelope and
TITLE OF INVENTION: Protein Allergens of the Species Cynodon
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,507
FILING DATE: 15-May-1995
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/969,875
FILING DATE: 10-October-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 1M1-049DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-441-507-17

```

Query Match 1.08; Score 16; DB 4; Length 594;
Best Local Similarity 100.0%; Pred. No. 26-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 319 TCTTGCCCATGGCGGC 334
|||||
DB 37 TCTTGCCCATGGCGGC 22

RESULT 48
US-07-969-875A-17/c
Sequence 17, Application US/07969875A
Patent No. 6441157
GENERAL INFORMATION:
APPLICANT: Singh, Mohan Hir;
APPLICANT: Smith, Penelope; and
APPLICANT: Knox, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation
STREET: 610 Lincoln Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/969,875A
FILING DATE: 30-October-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vanshstone, Darlene A.
REGISTRATION NUMBER: 35,279
REFERENCE/DOCKET NUMBER: 041.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6010
INFORMATION FOR SEQ. ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-07-969-875A-17

Query Match 1.08; Score 16; DB 4; Length 594;
Best Local Similarity 100.0%; Pred. No. 26-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 TCTTGCCCATGGCGGC 334
|||||
DB 37 TCTTGCCCATGGCGGC 22

RESULT 49
US-08-441-507-1/c
Sequence 1, Application US/08441507
Patent No. 6214358
GENERAL INFORMATION:
APPLICANT: Singh, Mohan Hir;
APPLICANT: Smith, Penelope; and
APPLICANT: Knox, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species Cynodon
NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:
ADDRESSEE: LAMIVE & COCKFIELD LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,507
FILING DATE: 15-May-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,875
FILING DATE: 30-October-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 1M1-0494V
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 662 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..435
US-08-441-507-1

Query Match 1.08; Score 16; DB 4; Length 662;
Best Local Similarity 100.0%; Pred. No. 2,10-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0;

QY 319 TCTTGCCCATGGCGGC 334
|||||
DB 58 TCTTGCCCATGGCGGC 43

RESULT 40
US-07-969-875A-1/c
Sequence 1, Application US/07969875A
Patent No. 6441157
GENERAL INFORMATION:
APPLICANT: Singh, Mohan Hir;
APPLICANT: Smith, Penelope; and
APPLICANT: Knox, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation
STREET: 610 Lincoln Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/969,875A
FILING DATE: 30-October-1992

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vanstone, Darlene A.
REGISTRATION NUMBER: 45,279
REFERENCE/DOCKET NUMBER: 041.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6010
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 662 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..435
US-07-969-875A-1

```

```

Query Match: 1.0%; Score 16; DB 4; Length 662;
Best Local Similarity: 100.0%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 319 TCTTGGCATGCGCGC 334
|||||
FH 58 TCTTGGCATGCGCGC 43

```

```

RESULT 41
US-09-288-143-11
Sequence 11, Application US/09288143
Patent No. 6433139
GENERAL INFORMATION:
APPLICANT: Brewer et al.
TITLE OF INVENTION: 53 Human Secreted Proteins
FILE REFERENCE: P2018P1
CURRENT APPLICATION NUMBER: US/99/288,143
CURRENT FILING DATE: 1999-04-08
EARLIER APPLICATION NUMBER: PCT/US98/21142
EARLIER FILING DATE: 1998-10-08
EARLIER APPLICATION NUMBER: 60/061,463
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,529
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/071,498
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,527
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,536
EARLIER FILING DATE: 1997-10-09
EARLIER APPLICATION NUMBER: 60/061,542
EARLIER FILING DATE: 1997-10-09
NUMBER OF SEQ ID NOS: 219
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 11
LENGTH: 680
TYPE: DNA
ORGANISM: Homo sapiens
US-09-288-143-11

```

```

Query Match: 1.0%; Score 16; DB 4; Length 680;
Best Local Similarity: 100.0%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1527 TAACTGCAAAAAAAA 1542
|||||
DB 640 TAACTGCAAAAAAAA 655

```

```

RESULT 42
US-09-198-603C-26
Sequence 26, Application US/09198603C

```

```

Patent No. 6347194
GENERAL INFORMATION:
APPLICANT: TULLY, Raymond E.
APPLICANT: CALTAGIRONE, G. Thomas
APPLICANT: MOYER, Shawn S.
APPLICANT: KONNING, Michael L.
TITLE OF INVENTION: EXPRESSION OF MANNOSYL-BINDING PROTEIN IN METHYLOTROPHIC YEAST
FILE REFERENCE: A7290
CURRENT APPLICATION NUMBER: US/98/603,603C
CURRENT FILING DATE: 1998-11-24
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 26
LENGTH: 714
TYPE: DNA
ORGANISM: CHICKEN
US-09-198-603C-26

```

```

Query Match: 1.0%; Score 16; DB 4; Length 714;
Best Local Similarity: 100.0%; Pred. No. 2.1e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1139 AATGAATCTGTGGCG 1154
|||||
DB 628 AATGAATCTGTGGCG 643

```

```

RESULT 43
US-08-441-507-20/C
Sequence 20, Application US/08441507
Patent No. 6214358
GENERAL INFORMATION:
APPLICANT: Singh, Mohan Bitt
APPLICANT: Smith, Penelope; and
APPLICANT: Knux, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species Cynodon
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,507
FILING DATE: 15-May-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,875
FILING DATE: 30-October-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 1M1-0490V
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

```


NAME/KEY: CDS
LOCATION: 1..738
FEATURE:
NAME/KEY: CDS
LOCATION: 742 759
US-08-441-507-20

Query Match 1.08; Score 16; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 2,1e-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 TCTGGCCATGGCGC 334
|||||
DB 355 TCTGGCCATGGCGC 340

RESULT 44

US-07-969-875A-20/C
Sequence 20, Application US/07969875A
Patent No. 6441157
GENERAL INFORMATION:
APPLICANT: Singh, Mohan Brij
APPLICANT: Smith, Penelope; and
APPLICANT: Knox, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation
STREET: 610 Lincoln Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/969,875A
FILING DATE: 30-October-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vanstone, Darlene A.;
REGISTRATION NUMBER: 35,279
REFERENCE/DOCKET NUMBER: 041.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6010
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 759 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..738
NAME/KEY: CDS
LOCATION: 742..759
US-07-969-875A-20

Query Match 1.08; Score 16; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 2,1e-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 319 TCTGGCCATGGCGC 334
|||||
DB 355 TCTGGCCATGGCGC 340

RESULT 45
US-08-441-507-3/C
Sequence 3, Application US/08441507
Patent No. 6214358
GENERAL INFORMATION:
APPLICANT: Singh, Mohan Brij
APPLICANT: Smith, Penelope; and
APPLICANT: Knox, Robert Bruce
TITLE OF INVENTION: Protein Allergens of the Species Cy
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,507
FILING DATE: 15-May-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/969,875
FILING DATE: 30-October-1992
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: 1M1-0490V
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 775 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..600
US-08-441-507-3

Query Match 1.08; Score 16; DB 4; Length 775;
Best Local Similarity 100.0%; Pred. No. 2,1e-02;
Matches 16; Conservative 0; Mismatches 0; Indels 0;

QY 319 TCTGGCCATGGCGC 334
|||||
DB 217 TCTGGCCATGGCGC 202

Search completed: December 7, 2002, 08:48:07
Job time : 649 secs

100

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 13:53:37 Search time 2254 Seconds

(without alignments)
2335.197 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325

Sequence: 1 MSDSKEPRVQGLGGLCCLG...

Scoring table:

OLIGO	60.0	Xgapex	60.0
Xgapop	60.0	Xgapex	60.0
Ygapop	60.0	Ygapex	60.0
Delop	6.0	Delop	7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 5

Total number of hits satisfying chosen parameters: 2258837

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL-frame-p2n.model -DBV-X1h
-O/cqn2_170SPFO_Spool/US0831458/runtat_05122002_104421_169377/APF_query_fasta_1.519
-DB-EST-QFMT-Iastap -SUFFIX-oli.rst -MINMATCH=0.1 -LOCBL=0 -LOCPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX=oliqo -TRANS-human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=5 -ALIGN=45 -MOP=LOCAL -OUTPMT=PTO
-NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-NO_XLPEX -NO_MMAP -LARGEOUTER -NFG_SCORES=0 -WAIT -IONOPLOP -FEV_TIMEOUT=120
-WANT_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DPLEXT=7

Database:

EST: *

1:	em_estba:
2:	em_esthum:
3:	em_estinu:
4:	em_estinu:
5:	em_estrov:
6:	em_estip1:
7:	em_estiro:
8:	em_hic:
9:	qb_est1:
10:	qb_est2:
11:	qb_hic:
12:	qb_est3:
13:	qb_est4:
14:	qb_est5:
15:	em_estfun:
16:	em_estom:
17:	qb_gss:
18:	em_gss_hum:
19:	em_gss_inu:
20:	em_gss_plu:
21:	em_gss_vrt:
22:	em_gss_fun:
23:	em_gss_hum:
24:	em_gss_mus:
25:	em_gss_other:
26:	em_gss_pro:
27:	em_gss_rod:

Pred. No. is the number of results predicted by chance 1
score greater than or equal to the score of the result
and is derived by analysis of the total score distribut

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Loc
1	153	47.1	762	14	B1828875
2	153	47.1	762	9	AL584658
3	128	39.4	706	10	AV645755
4	125	38.5	786	13	B1827955
5	125	38.5	828	10	BEC62903
6	107	32.9	503	14	R98113
7	97	29.8	722	9	AU140165
8	60	18.5	419	14	H90968
9	50	15.4	486	14	R80140
10	47	14.5	833	9	AL570816
11	45	13.8	359	10	AV645629
12	39	12.0	411	9	AA448145
13	19	5.8	453	14	R80141
14	15	4.6	499	9	AA441250
15	11	3.4	284	9	AL614060
16	11	3.4	388	9	AA510960
17	11	3.4	635	10	BH583702
18	10	3.1	666	14	H90460
19	10	3.1	666	17	AG139337
20	9	2.8	69	17	AC782151
21	9	2.8	230	9	AL508815
22	9	2.8	262	10	BH466869
23	9	2.8	288	17	BH52311
24	9	2.8	300	14	C57235
25	9	2.8	386	10	AME34972
26	9	2.8	461	12	B3554984
27	9	2.8	482	10	AM452345
28	9	2.8	489	9	AA388306
29	9	2.8	581	12	BH422184
30	9	2.8	607	12	BH041962
31	9	2.8	610	13	B1481901
32	9	2.8	614	10	BH637495
33	9	2.8	633	10	BH659390
34	9	2.8	719	14	BG574692
35	9	2.8	810	17	BH594642
36	9	2.8	811	17	CNS02X10
37	9	2.8	879	13	B1757640
38	9	2.8	897	17	CNS04Y80
39	9	2.8	936	17	CNS03H1K
40	9	2.8	977	17	CNS022K0
41	9	2.8	1018	17	B17648
42	9	2.8	1101	17	CNS0038S
43	9	2.8	3108	11	AF12661K
44	8	2.5	164	13	BH441797
45	8	2.5	171	14	BH71071

ALIGNMENTS

RESULT 1
B1828875
LOCUS
DEFINITION
603075032F1 NIH_MGCT_119 Homo sapiens cDNA clone
B1828875
ACCESSION
B1828875
VERSION
B1828875.1 GI:15949425
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota: Metazoa: Chordata: Granulata: Vertebrata
Mammalia: Eutheria: Primates: Catarrhini: Homini:
1 (bases 1 to 765)
REFERENCE
NIH-MGCT <http://mgc.nhl.nih.gov/>
AUTHORS
National Institutes of Health, Mammalian Gene Cla


```

DB 429 GCTGACGCTGATTCGCTAGAGAAATCTCAATCTTATAGATCTATCTAGTAACTTAA 488
OY 146 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 155
DB 489 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 548
OY 156 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 168
DB 549 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 587

RESULT 3
AV645755 706 bp mRNA linear EST 15 JAN 2002
LOCUS AV645755 GIC Homo sapiens cDNA clone GICAD05 3, mRNA sequence.
ACCESSION AV645755
VERSION AV645755.1 GI:9866769
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 706)
AUTHORS Xu, X., Huang, J., Xu, Z., Qian, B., Zhu, Z., Yan, Q., Cai, T., Zhang, X.,
Xiao, H., Qu, J., Liu, F., Huang, Q., Cheng, Z., Lin, P., Hu, W.,
Shen, K., Lu, G., Fu, G., Zhong, M., Xu, S., Gu, W., Huang, W., Zhao, X.,
Hu, G., Gu, J., Chen, Z., and Han, Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (25), 15084-15094 (2001)
21625106
CONTACT: Zequan Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801914 (ex 45)
Fax: 86-21-50801922
Email: hanzq@hgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Source
Location/Qualifiers
1..706
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GICAD05"
/clone_lib="GIC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 168 a 198 c 176 g 162 t 2 others
ORIGIN
XhoI"

Alignment Scores:
Pred. No.: 1 98e-122 length: 706
Score: 128.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.38% Indels: 0
DB: 10 gaps: 0

US-09-831-458a-12 (1-325) x AV645755 (1-706)
OY 188 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 207
DB 145 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 204
OY 208 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 227
DB 205 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 264
OY 228 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 247

```

```

DB 205 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 248
OY 248 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 325
DB 325 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 385
OY 385 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 445
DB 445 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 505
OY 505 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528

RESULT 4
B1827055 786 bp mRNA linear EST 15 JAN 2002
LOCUS B1827055 GIC Homo sapiens cDNA clone 1, mRNA sequence.
ACCESSION B1827055
VERSION B1827055.1 GI:15948605
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 786)
AUTHORS NIH-MGC http://www.ncbi.nlm.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: greg@lmi.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution info:
http://image.llnl.gov
Plate: L14M1416 row: 4 column: 08
High quality sequence stop: 776.

FEATURES
Source
Location/Qualifiers
1..786
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5167447"
/clone_lib="NIH-MGC_119"
/tissue_type="medulla"
/lab_host="pH10P"
/note="Organ: brain; Vector: pCMV-Script;
Site_2: EcoRV (destroyed); RNA source:
anonymous male age 27. Library is ori-
directionally cloned (EcoRV site is des-
troyed). Average insert size 1.3 kb,
0.9-3 kb. Library is normalized and
full-length clones and was constructed
(invitrogen). Research Genetics Trace
this is a NIH-MGC library."
BASE COUNT 212 a 193 c 241 g 140 t
ORIGIN

Alignment Scores:
Pred. No.: 3e-119 length: 786
Score: 125.00 Matches: 125
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.46% Indels: 0
DB: 13 gaps: 0

```


[illegible]

REFERENCE AUTHORS	TITLE	JOURNAL MEDLINE COMMENT
Xu, X., Huang, J., Xu, Z., Guan, R., Zhu, Z., Yan, G., et al.	(bases 1 to 359)	
Xiao, H., Ou, J., Lin, F., Huang, Q., Cheng, Z., Li, N.		
Shen, K., Lu, G., Chen, Z., Zhong, M., Xu, S., Gu, W., Hu,		
Hu, G., Gu, D., Fan, Z., and Han, Z.		
	Insight into hepatocellular carcinogenesis at the	
	by comparing gene expression profiles of hepatocarcinoma	
	with those of corresponding noncancerous liver	
	Proc. Natl. Acad. Sci. U.S.A. 96 (26), 15089-1509	
	21625106	
	Contact: Zequan Han	
	Chinese National Human Genome Center at Shanghai	
	351 Guo Shoujing Road, Shanghai Hi-Tech Park,	
	201203, P. R. China	
	Tel.: 86-21-50801919 (x. 45)	
	Fax: 86-21-50801922	
	Email: hanzq@hgc.sh.cn	
	This clone is available at CHGC in Shanghai.	
	Location/Qualifiers	
	1. 359	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone_lib "GICAC012"	
	/clone_lib "GIC"	
	/library_type="Corresponding non cancerous"	
	/dev_stage="Adult"	
	/lab_host="SOLR"	
	/note="Vector: phuescript sk(-); Site: XhoI"	
BASE COUNT	85 a 111 c 91 g 72 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	1,72e-36	Length: 359
Score:	45.00	Matches: 45
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	13,858	Indels: 0
DB:	10	Gaps: 0
US-09-831-458A-12 (1-325) x AV45629 (1-459)		
OY 188	GI:4949CysAR9HScysSProlysAspRPRhphepGlnGlyAs-	
DB 145	GAAAGCTGTGGCGCCACACTGTGACCAAGAGTGGAAATTCTTCCAAAGGAAAT	
OY 208	MetSerAnsSerGlnArqAsnTPHtSAspSeValTPHrAlaCysGlnGln-	
DB 205	ATGCTAACCTCCAGCGGAACTGGTACACTCGTCAACCGCGCTGTCAGAA	
OY 228	GTCLeuValValIle 232	
DB 265	CAGCTCCGTCGTAATC 279	
RESULT 12		
LOCUS	AA448145	411 bp
DEFINITION	5' mRNA sequence: c13315_2H1 Homo sapiens cDNA	
ACCESSION	AA448145	
VERSION	AA448145.1	GI:2161415
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Ekabof, M., Melzer, C., Hatakeyama, T., Verstra-	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homin-	
	1 (bases 1 to 411)	
	Hillier, L., Allen, M., Bowles, J., Dubouche, J., Fe-	
	kiach, A., Lacey, M., Le, N., Lennon, G., Marra, M., F-	
	Schellenberg, K., Stepien, M., Tan, F., Theisnig,	
	J., Waterston, R., and Wilson, R.	
TITLE	WashU-Merck EST Project 1997	
JOURNAL	Unpublished (1997)	

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
 1 (bases 1 to 381)
 HILLIER, L., CLARK, N., DUBUQUE, T., ELLISTON, G., HAWKINS, M., HOLMAN,
 M., HULMAN, M., KICABA, T., LE, M., LENNON, G., MARRAS, M., PARSONS, J.,
 RIKLIN, L., ROHLING, T., SOARES, M., TAN, F., TROVASKIS, F., WATERSTON,
 R., WILLIAMSON, A., WOHLDMANN, P. and WILSON, R.
 The WashU-Merck EST Project
 TITLE Unpublished (1995)
 JOURNAL Contact: Wilton RK
 COMMENT Washington University School of Medicine
 4444 Forest Park Parkway, Box 8001, St. Louis, MO 63110
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Insert Size: 712
 High quality sequence stops: 331
 Source: IMAGE Consortium, LIND
 This clone is available royalty-free through LIND; contact the
 IMAGE Consortium (info@image.lind.gov) for further information.
 Insert Length: 712 Std Error: 0.90
 Seq primer: M13RPI
 High quality sequence stop: 331.
 Location/Qualifiers
 1..381
 /organism="Homo sapiens"
 /db_xref="GDB:3790505"
 /db_xref="taxon:9606"
 /clone="IMAGE:241372"
 /clone_lib="Soares fetal liver spleen INF15"
 /sex="male"
 /dev_stage="20 week post conception fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ. Liver and Spleen; Vector: p1713b (Pharmacia)
 with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
 1st strand cDNA was primed with a Pac I - oligo(dT) primer
 15' AACTCGACGATTAATTAAGATCTTTTCTTTTCTTTT 3';
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Pac I and cloned into the Pac I
 and Eco RI sites of the modified p1713 vector. Library
 went through one round of normalization. Library
 constructed by Benito Soares and M. Fatima Boudado."
 BASE COUNT 68 a 97 c 88 g 125 t 3 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.7 length: 381
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Host Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.08% Indels: 0
 DB: 14 gaps: 0
 US-09-831-458a-12 (1-325) x H90360 (1-381)
 QY 316 LysLysProAlaAlaGlySpharGASpGlu 325
 |||||
 DB 2 AAAAAGCCGCGACGCTGCTTCAGACGAA 31
 RESULT 19
 AC141937/c AG143937 666 bp DNA linear GSS:08-JAN-2002
 LOCUS Pan troglodytes RNA, clone: RP43-004M13.17, genomic survey
 DEFINITION
 sequence:
 AC143937
 VERSION AG143937.1 GI:16673615
 KEYWORDS GSS.
 SOURCE Pan troglodytes male lymphocytes DNA, clone: RP43-004M13.17.
 ORGANISM Male BAC Library clone: RP43-004M13.17.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.

REFERENCE 1
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, I. D.,
 Toki, Y., Watanabe, H. and Sakaki, Y.
 TITLE BAC end sequences of Library RP43-43
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 666)
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, I. D.,
 Toki, Y., Watanabe, H. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02 AUG 2001) Asao, Fujiyama, The Institute
 and Chemical Research (RIKEN), Genomic Sciences
 1-7-22 Sushiro-cho, Tsukuba, Ibaraki, Yokohama, Kanagawa,
 (E-mail: chimp@rsc.riken.go.jp, Tel: 81-45-503-9171,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library
 end was generated during the R&D process and may
 of clone tracking errors.
 PRIMER Sequencing: 17
 LIBRARY
 Vector : pBR322
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.
 Location/Qualifiers
 1..666
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="RP43-004M13.17"
 /sex="male"
 /cell_type="lymphocytes"
 /clone_lib="Rpi-43 Chimpanzee Male BAC"
 BASE COUNT 121 a 270 c 147 g 126 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 9.75 length: 666
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Host Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.08% Indels: 0
 DB: 17 gaps: 0
 US-09-831-458a-12 (1-325) x AG143937 (1-666)
 QY 270 GlySerProLeuSerProSerPheGlnArg 279
 |||||
 DB 320 GGTTCCTCTCTGCGCCGATTCGACGCG 291
 RESULT 20
 AC2782451 649 bp DNA linear
 LOCUS 2M002211R Mouse 1995 Flast13 ung21M library Mus
 DEFINITION
 sequence:
 AC2782451
 VERSION AZ782451 1 GI:12916186
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
 1 (bases 1 to 649)
 Dunn, D., Aoyagi, A., Bartol, M., Beatty, T., Brown,
 Islam, H., Longacre, S., Mahmood, M., Meegen, E.,
 M., Rose, M., Rose, B., Sakai, F., Tinney, A.,
 and Wright, D., Weiss, R.
 TITLE Mouse whole genome scaffolding with paired end
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Km. 308, Biomedical Polymers Research Bldg., 20
 84112, USA

DEFINITION C57235 Yui Kohara unpublished cDNA *Caenorhabditis elegans* cDNA
 clone yk2752 3', mRNA sequence.

ACCESSION C57235
 VERSION 057235.1 GI:2415266
 KEYWORDS EST.
 SOURCE *Caenorhabditis elegans*.
 ORGANISM *Caenorhabditis elegans*.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae
 1. Rhabditidae; Rhabditinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 300)
 AUTHORS Kohara, Y., Motobashi, T., Tabara, H., Watanabe, H., Sugimoto, A., Sano, M., Miyata, A. and Nishizaki, A.
 TITLE Expression map of the *C. elegans* genome
 JOURNAL Unpublished (1996)
 COMMENT Contact: Yui Kohara
 Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykoha@lab.nig.ac.jp.
 Location/Qualifiers
 1..300
 /organism="Caenorhabditis elegans"
 /strain="YK1489 him-8(e1489)"
 /db_xref="taxon:6239"
 /clone="YK2752"
 /clone_lib="Yui Kohara unpublished cDNA"
 /sex="hermaphrodite, male"
 /tissue="type="whole animal"
 /dev_stage="varied"

BASE COUNT 63 a 81 c 72 g 82 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 37.7 Length: 300
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 DE: 14 Gaps: 0

US-09-831-458a-12 (1-325) x C57235 (1-300)

QY 280 TTTTTPASerGlyClnProAsn 288
 ||||||||||||||||||||
 Db 110 TACTGSAATTCCGAGAGCCCAATAT 84

RESULT 25
 AM534972 386 bp mRNA linear EST 06-MAR-2000
 LO 705
 DEFINITION U1-R-BS0-any-f-08-0-01 s1 ut-p-R50 Ratios portegious cDNA clone
 ACCESSION AM534972
 VERSION AM534972.1 GI:7177386
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM *Rattus norvegicus*.
 Eukaryota; Metazoa; Chordata; Cladida; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 386)
 REFERENCE 1 (bases 1 to 386)
 AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genoma Res 6 (9): 791-806 (1996)
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565

Email: msosares@iue.wisc.iowa.edu
 The sequence contained an oligo dt track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NotI site
 and the oligo-dT track served to verify it as a clone from the
 non-normalized embryo at 13 dpc library cDNA library preparation:
 M.B. Soares Lab clone distribution: clones will be available
 through Research Genetics (www.resgen.com) the following repetitive
 elements were found in this cDNA sequence: 1-22,
 >AL1_F1c1aw_complexity 23-151, >BC1_MMScRNA 211-294,
 >SINE1/SINE/84(B5)
 Seq primer: M13 Forward
 polyA-Yes.

FEATURES
 Source
 1..386
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="U1-R-BS0-any-f-08-0-01"
 /clone_lib="U1-R-BS0"
 /dev_stage="embryonic 13 dpc"
 /lab_host="DHIb (Life Technologies)"
 /note="Vector: pTZ130-pac (Pharmacia) with a modified
 polylinker site. 1..Not I; Site 2.. Eco P1; The U1-R-BS0
 library is derived from 13 dpc whole embryo tissue. For a
 detailed description of the library from which this clone
 was derived, please visit our web site at
 ralest.eng.iowa.edu.
 TAG_LIB=U1-R-BS0
 TAG_TISSUE=embryo at 13 dpc
 TAG_SEQ=AAATCC"

BASE COUNT 80 a 88 c 68 g 150 t
 ORIGIN

Alignment Scores:
 Pred. No.: 52.4 Length: 386
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 DE: 10 Gaps: 0

US-09-831-458a-12 (1-325) x AM54972 (1-386)

QY 224 GTCVATATGACACACTTGTCATCATT 250
 ||||||||||||||||||||
 Db 224 GAGCTCATGACACACTTGTCATCATT 250

RESULT 26
 BG554984 461 bp mRNA linear EST 09-APR-2001
 LO 705
 DEFINITION BG554984
 VERSION BG554984
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM *Xenopus laevis*.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Anphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 461)
 REFERENCE 1 (bases 1 to 461)
 AUTHORS Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Page, D.,
 Martin, J., Wylie, F., Underwood, K., Theising, B., Bowers, Y., Person,
 B., Gibbons, M., Harcey, N., Ritter, E., Jackson, Y., McCann, R.,
 Waterston, R. and Wilson, R.
 TITLE WASHU Xenopus EST Project, 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: Sandy Clifton, Ph.D.
 WASHU Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
Library constructed by Life Technologies. DNA Sequencing by:
Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: InfoImage.Llnl.gov
Seq primer: -40bp from ribco
High quality sequence stop: 439
Location/Qualifiers

FEATURES

source

1. 461
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_image="408871"
/clone_id="NICHD XGC He1"
/lab_host="DH10B (phage-resistant)"
/note="Organ: heart; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI. Cloned individually. Primer oligo dt.
Average insert size 1.6 kb. Constructed by life
Technologies." 141 a 89 c 81 g 150 t

BASE COUNT

ORIGIN

Alignment Scores:
Pred. No.: 66.1 Length: 461
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x BG554984 (1-461)

UY 217 ASPserValThrAlaGlyClnGluVal 225

Db 19 GATTTCAGTACAGCTGCCAGGAAGTA 45

RESULT 27

AA452345/C

LOCUS 482 bp mRNA linear EST 17-FEB-2000

DEFINITION IMAGE3068165 3, mRNA sequence.

ACCESSION AA452345

VERSION AA452345.1 GI:6993121

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 482)

AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Journal Comment

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: rga@nci.nih.gov

Oligo-dt track not found. Not 1 site shown in beginning of sequence

is likely internal to the message. cDNA library preparation: M.B.

Sources Lab clone distribution: NCI-CCAP clone distribution

Information can be found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bdrrp/image/image.html

Seq primer: M13 Forward

POLYA-NO.

Location/Qualifiers

1. 482

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="3068165"

/clone_id="NICL_CGAP_Subs"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker. Site_1: Not I; Site_2: Eco RI; NCI-CCAP_Subs

is a subtracted library derived from NCI-CCAP_Sub4. The

NCI-CCAP_Subs library had 3 million recs

single-stranded DNA preparation of NCI-CCAP
as a tracer in a subtractive hybridization
experiment. The IMAGE pool (NCI-CCAP_Kid
3334-3347, 4682-4687, 4798-4803 (IMAGE)
1323376-1323915, 1456008-1456775, 1506755-
1507492 (IMAGE) K145 pool 1 LLAM 3348-3349, 4742-
4743 (IMAGE) Clones 1323912-1325831, 1471866-
1492104 (1492255); NCI-CCAP_Lib pool 1 Lib
3851-3854 (IMAGE) Clones 1414920-1417555
(NCI-CCAP_Lib pool 1 LLAM 3164-3167,
3733-3735 (IMAGE) Clones 1257096-125786,
1475592-1475743); NCI-CCAP_P12 pool 1
2758-2759, 4062-4068 (IMAGE) Clones 94,
1101192-1101459, 1217928-1220615); NC
LLAM 2644-2645, 2871-2872 (IMAGE) Clones
1144584-1144585 (10% of the driver p
pool of 3,410, 11,499, 11,499, 11,499, 11,499,
Clones 2708616-2710545) and NCI-CCAP
Clones 2710546-2712455 (10% of the
) plus a pool of 11,136 clones from NCI-CCAP
Clones 2712456-2733541 (10% of the driver
plus a pool of 5,472 clones from NCI-CCAP
Clones 2733542-2738949) (70% of the driver
Subtraction was performed as previously
Lennon & Soares (1996): Normalization
Two Approaches to Facilitate Gene Discovery
Research 6, 791-806.
TAG_Lib=NCI-CCAP_Lib
TAG_Tissue=germ cell
TAG_SEQ=AAATTC

BASE COUNT 91 a 135 c 157 g 98 t 1 other

ORIGIN

Alignment Scores:

Pred. No.: 70.1 Length: 482
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x AA452345 (1-482)

UY 34 AlaGlyValLeuValAlaIleGluVal 42

Db 425 GCTGGAGTCTGGTGGCCATCTG51G 999

RESULT 28

AA388306/C

LOCUS 499 bp mRNA linear

DEFINITION IMAGE:775542 5', mRNA sequence.

ACCESSION AA388306

VERSION AA388306.1 GI:2011272

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 499)

AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., D'Arcy,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J.,
Schellenberg, K., Strohman, M., Tan, F., Underwood, J.,
Theising, B., Wyllie, J., Lennon, G., Soares, R., W
Waterson, R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Maria M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO

Tel: 314 286 1800


```

seedling"
/lab_host="DH10B"
/Note="Vector: pBluescript 11 SK+, Site1: EcoRI; Site2:
XhoI. The cDNA library was constructed from mRNA isolated
from degenerating ocytodes of 2 week old seedlings from
p1468916. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion.
The cDNA fragments were directionally cloned into the
EcoRI XhoI restriction site of the pBluescript vector. The
ligated cDNA fragments were transformed into DH10B host
cells (GibcoBRL). This library was constructed in the
laboratory of Dr. Randy Shoemaker."
BASE COUNT      111 a      179 c      160 g      157 t
ORIGIN
Alignment Scores:
Pred. No.:      94.7      Length:      607
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels:      0
DB:              12      Gaps:      0
US-09-831-458a-12 (1-325) x BG041052 (1-607)
OY      239 AsnPhelenglnleuclntrseratg 247
Db      70 AACCTCCTCAGCTTCAACCTCTCCG 96
RESULT 11
LOCUS      B1481901      610 bp      mRNA      linear      EST 28-AUG-2001
DEFINITION Drosophila melanogaster normalised Embryo pflc-1
ACCESSION   B1481901
VERSION     B1481901.1 GI:15321110
KEYWORDS    EST
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 610)
AUTHORS     Stapleton, M., Brockstein, P., Hong, L., Tyler, D., Berman, B., Carlson
, R., Champe, M., Chavez, C., Dorsett, V., Farran, D., Fris, E., George
, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Li, G., Mista, S.,
Mungall, C. J., Nuppon, T., Parle, T., Paragas, V., Park, S.,
Phouanavong, S., Wan, K., Yu, G., Lewis, S. E., Celisner, S. and Rubin
, G. M.
JOURNAL     BDCP/HMI RE Drosophila EST project
COMMENT     Unpublished (2001)
CONTACT     Stapleton, M.
BOGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estafruitfly_berkeley.edu
hit genomic AEU03489; arm: d1608641.9111671
estimated-cyto: 6178-61E1: 05/16/2001
Plate: RF 645 row: B column: 6
High quality sequence stop: 540.
location/Qualifiers
1..610
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="RE64518"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pflc-1"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"

```

```

/Note="Organ: embryo; Vector: pflc1; Site1:
BamHI; Site2: XhoI. The library was constructed from
the RIKEN cDNA library. The library was normalized and
cre recombinase. Plasmid cDNA library."
BASE COUNT      202 a      117 c      146 g      145 t
ORIGIN
Alignment Scores:
Pred. No.:      95.3      Length:      610
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels:      0
DB:              13      Gaps:      0
US-09-831-458a-12 (1-325) x B1481901 (1-610)
OY      272 ProLuserProserPheclArqTyr 280
Db      260 CCACCTTCCTCCCTCCCTCCACAGTAT 264
RESULT 32
LOCUS      B637495      614 bp      mRNA      linear
DEFINITION B637495 RIKEN full-length enriched, adult male
musculus cDNA clone A30080821 5', mRNA sequence.
ACCESSION   B637495
VERSION     B637495.1 GI:16473280
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata;
Mammalia; Eutheria; Rodentia; Sclerotiniath; Murinae;
1 (bases 1 to 614)
REFERENCE   1
AUTHORS     Arakawa, T., Carninci, P., Furuta, S., Furuo, M., Hara, A.,
Hiramoto, K., Hori, P., Ishii, Y., Ito, M., Kawai,
, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, R.,
Okazaki, Y., Okada, T., Saito, P., Sakai, C., Sakai,
, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogai,
, Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y.,
Muramatsu, M. and Hayashizaki, Y.
JOURNAL     RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
COMMENT     Unpublished (2001)
CONTACT     Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group
Science Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research
1-7-22 Suehiro cho, Tsukuba, Ibaraki, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome-res.riken.go.jp/
Carninci, P., Shibata, Y., Hayashizaki, N., Sugahara, Y.,
, M., Kondo, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper seq
prepare full-length cDNA libraries for rapid disc
genes. Genome Res. 10 (10), 1617-1630 (2000)
Wang, K., Fujitake, S., Inoue, K., Iiyazawa, Y., Iizawa
, M., Kawai, J., Kikuchi, Y., Kikuchi, Y., Kikuchi, Y.,
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y.,
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system
sequencing pipeline with 484 multiplexed seqs
10 (11), 1757-1771 (2000)
Kondo, H., Hayashizaki, Y., Shibata, K., Itoh, M.,
, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length
encyclopedia: real-time sequence clustering for
nonredundant cDNA library. Genome Res. 11 (2),
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H.,
, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J.,
Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDN

```



```

DB 313 TACCAGACTGACCGCAGTGAAGCT 339
RESULT 34
B0574692/c 719 bp mRNA linear EST 19-JUN-2002
LOCUS B0574692
DEFINITION B0574692: hsa-b-14-0-01 s1 NCL-GAP-Ch2 Homo sapiens cDNA clone
UT-H-E21-bba-h-14-0-01 3, mRNA sequence.
ACCESSION B0574692
VERSION B0574692.1 GI:21478039
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates, Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 719)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rga@pilot.fda.gov
Tissue Procurement: Dr. Steven Gitlin/ Rush Presbyterian, Dept. of
Orthopedics
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 542-587, SAT-rich; low complexity 599-719, ACU
Seq primer: M13 FORWARD
POLY(A)=Yes

FEATURES
Source
Location/Qualifiers
1..719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="01-H-E21-bba-h-14-0-01"
/clone_lib="NCL-GAP-Ch2"
/tissue_type="Chondrosarcoma Grade II"
/dev_stage="Adult"
/lab_host="DH10b (Life Technologies)"
/note="Organ: Left Pelvis; Vector: pT73-Pac (Pharmacia)
with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCL-GAP-Ch2 is a normalized cDNA library containing the
following tissue(s): Chondrosarcoma Grade II. The library
was constructed according to Bonaldi, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
TCATCAGCT.
TAG_LIB="01-H-E21"
TAG_TISSUE="grade-2-chondrosarcoma"
TAG_SEQ="ATCTAATG"

BASE COUNT 202 a 158 c 150 g 208 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 118 Length: 719
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
Gaps: 0
DB: 14 Gaps: 0

US-09-831-458a-12 (1-325) x B0574692 (1-719)
UY 44 A1aC1yVAlleuVAla1a1leuVal 42
|||||

```

```

DB 424 CCGAGCTGTTCGCGCACTTCATC 338
RESULT 35
B0694642
LOCUS B0694642 810 bp DNA linear EST 19-JUN-2002
DEFINITION B0694642: Bos_2_Ka Brassica oleracea genomic DNA
sequence.
ACCESSION B0694642
VERSION B0694642.1 GI:18765222
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core
Rosidae; eurosids II; Brassicales; Brassicaceae;
1 (bases 1 to 810)
AUTHORS Town, C.D., Van Arck, S., Utrecht, T. and Fraser, J.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Other_GSSs: B069361F
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, U.S.A.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Ost
Seq primer: TR
Class: sheared ends.

FEATURES
Source
Location/Qualifiers
1..810
/organism="Brassica oleracea"
/strain="T01000B3"
/db_xref="taxon:4712"
/clone="B069361F"
/clone_lib="B069361F"
/note="Vector: pHost; Site 1: BstXI; 2:
genomic DNA inserted into pHost using Bst
X1"

BASE COUNT 292 a 96 c 140 g 282 t
ORIGIN
Alignment Scores:
Pred. No.: 138 Length: 810
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
Gaps: 0
DB: 17 Gaps: 0

US-09-831-458a-12 (1-325) x B0694642 (1-810)
UY 73 G1yG1u1euSerg1u1ySerg1y1eu 81
|||||
DB 758 GGAGCACTTCCGAGAAATCAAAAT 784

RESULT 36
LOCUS C0502XJ0/c 811 bp DNA linear EST 19-JUN-2002
DEFINITION C0502XJ0: Tetradon nigrovittatus genome survey sequence 17810 of library G from Tetradon nigrovittatus
sequence.
ACCESSION AL218447
VERSION AL218447.1 GI:7877256
KEYWORDS GSS, genome survey sequence.
SOURCE Tetradon nigrovittatus.
ORGANISM Tetradon nigrovittatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodonidae; Tetradon.
1 (bases 1 to 811)
AUTHORS Frost-Croall, J.H., Hall, J.W., Dastgheib, C., Mowbray, J.,
Bernot, A., Fitzmales, C., Wilcock, P., Brothier, P.,

```

TITLE Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 811)
 AUTHORS Forest-Collins,H., Billon,C., Pasiva,C., Fitzames,C., Fischer,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Herron,A. and
 Weissbach,J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 811)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 FEATURES
 source Location/Qualifiers
 1..811
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_11b="g"
 /note="Genoscope sequence ID : COAG178D005LP1-end : 77"
 BASE COUNT 213 a 193 c 199 g 206 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 138 Length: 811
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conserved: 9
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 Gaps: 0
 US-09-831-458a-12 (1-325) x CNS02XJQ (1-811)
 CY 22 AAAAAAAAAAAAAAAAAAAAAA 30
 DB 764 GCGTCCTCTCTGCACTCTCTTCTT 738
 RESULT 37 879 bp mRNA linear EST 25-SEP-2001
 ID US B1757640/c 603027936F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196265.5,
 DEFINITION mRNA sequence.
 ACCESSION B1757640
 VERSION B1757640
 KEYWORDS EST.
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 879)
 NIH-MGC http://mgi.mcg.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: rgs@bimail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I M A G E Consortium (LMN)
 Clone distribution: MGC clone distribution information can be
 found through the I M A G E Consortium/IMN at:
 http://image.llnl.gov
 Plate: LLAM1496 row: h column: 10
 High quality sequence stop: 410.
 Location/Qualifiers
 1..879
 /organism="Homo sapiens"

/db_xref="taxon:9606"
 /clone_11b="MA3E:5198265"
 /clone_11b="NIH_MGC_114"
 /lab_host="DH10B"
 /notes="Organic brain; Vector: pCMV-Sport6; Site:1; Notif:
 Site:2; Esasy (b-stringed); DNA source: anonymous pool of 6
 male brains, age range 23-27 yrs. Library is oligo-dT
 primed and directionally cloned (kcoov site is destroyed
 upon cloning). Average insert size 1.5 kb. Insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code: 019. Note:
 this is a NIH-MGC Library."
 BASE COUNT 238 a 227 c 253 g 161 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 154 Length: 879
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conserved: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 DB: 13 Gaps: 0
 US-09-831-458a-12 (1-325) x B1757640 (1-879)
 CY 16 GTCGTCCTCTCTGCACTCTCTTCTT 24
 DB 137 GCGTCCTCTCTGCACTCTCTTCTT 111
 RESULT 38 897 bp DNA linear CSS 26 JUL 2000
 ID US B1757640/c 603027936F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196265.5,
 DEFINITION mRNA sequence.
 ACCESSION B1757640
 VERSION B1757640
 KEYWORDS EST.
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthopterygii; Acanthopterygii; Perciformes; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 897)
 Forest-Collins,H., Billon,C., Pasiva,C., Bouneau,L., Fischer,C.,
 Herron,A., Fitzames,C., Winkler,P., Brothier,P., Quetier,F.,
 Saurin,W. and Weissbach,J.
 Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 REFERENCE 2 (bases 1 to 897)
 AUTHORS Forest-Collins,H., Billon,C., Pasiva,C., Bouneau,L., Fischer,C.,
 Bouneau,L. and Weissbach,J.
 TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Unpublished (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 FEATURES
 source Location/Qualifiers
 1..897

```

/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="03701"
/clone_lib="A"
/notes="Genoscope sequence ID : C0AA037CE01A1 end : T3"
ORIGIN
BASE COUNT      249 a      147 c      146 g      263 t      62 others
ORIGIN

Alignment Scores:
Pred. No.:      158      Length:      897
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels:      0
DB:              17      Gaps:      0

US-09-831-458A-12 (1-325) x CNS04YKQ (1-897)

OY      27  LeuLeuSerPheMeLeuLeuAlaGly 35
        |||||
Db      207 CTCTTATCTTATCTCTCTCTCTCTCT 233

RESULT 39
CNS03HIK      936 bp      DNA      linear      GSS 17-MAY-2000
Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
027A01 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION      AL244325.1 GI:7965337
VERSION      GSS: genome survey sequence.
KEYWORDS      Tetraodon nigroviridis.
SOURCE      Tetraodon nigroviridis.
ORGANISM      Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 936)
AUTHORS      Roest-Crollius, H., Jallion, O., Dasilva, C., Bonneau, L., Fisher, C.,
Bernot, A., Fizames, C., Mincker, P., Brotlier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
HUMAN gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL      Unpublished
COMMENT      2 (bases 1 to 936)
AUTHORS      Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,
Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
JOURNAL      3 (bases 1 to 936)
AUTHORS      Genoscope.
TITLE      Direct Submission
COMMENT      Submitted (12-APR-2000)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
Source
1..936
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="027A01"
/clone_lib="G"
/notes="Genoscope sequence ID : C0R027AA01SP1 end :
PUC-ori"
BASE COUNT      276 a      178 c      175 g      301 t      6 others
ORIGIN
ORIGIN

Alignment Scores:
Pred. No.:      167      Length:      936
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels:      0
DB:              17      Gaps:      0

US-09-831-458A-12 (1-325) x CNS04YKQ (1-897)

OY      22  AlLeuValLeuGlnLeuSerPhe 40
        |||||
Db      295 GCTCTGCTCTCTCTCTCTCTCTCT 269

```

```

Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels: 0
DB:              17      Gaps: 0

US-09-831-458A-12 (1-325) x CNS03HIK (1-897)

OY      23  LeuValLeuGlnLeuSerPheMet 41
        |||||
Db      252 CTCGTTTACAGCTACTGCTCTTTATG 278

RESULT 40
CNS022KQ      977 bp      DNA      linear      GSS 17-MAY-2000
Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
227H14 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION      AL178281
VERSION      AL178281.1 GI:7816338
KEYWORDS      GSS: genome survey sequence.
SOURCE      Tetraodon nigroviridis.
ORGANISM      Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontidae; Tetraodon.
REFERENCE      1 (bases 1 to 977)
AUTHORS      Roest-Crollius, H., Jallion, O., Dasilva, C., Bonneau,
Bernot, A., Fizames, C., Mincker, P., Brotlier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.
HUMAN gene number estimate provided by genome wide
analysis using Tetraodon nigroviridis DNA sequence
JOURNAL      Unpublished
COMMENT      2 (bases 1 to 977)
AUTHORS      Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C.,
Bonneau, L., Billault, A., Quetier, F., Saurin, W.,
Weissenbach, J.
Characterization and repeat analysis of the compact
freshwater pufferfish Tetraodon nigroviridis
Unpublished
JOURNAL      3 (bases 1 to 977)
AUTHORS      Genoscope.
TITLE      Direct Submission
COMMENT      Submitted (12-APR-2000)
This sequence is a single read and was generated as
part of a large scale clone-end sequencing project of the Tetraodon
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
Source
1..977
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="227H14"
/clone_lib="G"
/notes="Genoscope sequence ID : C0AS227H00
PUC-ori"
BASE COUNT      211 a      295 c      282 g      182 t      7 others
ORIGIN
ORIGIN

Alignment Scores:
Pred. No.:      176      Length:      977
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      2.77%      Indels: 0
DB:              17      Gaps: 0

US-09-831-458A-12 (1-325) x CNS022KQ (1-977)

OY      22  AlLeuValLeuGlnLeuSerPhe 40
        |||||
Db      295 GCTCTGCTCTCTCTCTCTCTCT 269

```

RESULT 41
B12648 1018 bp DNA linear GSS 14 MAY 1997
DEFINITION B264-17.1 rat Arabidopsis thaliana genomic clone F26F4, DNA sequence.
ACCESSION B12648
VERSION B12648
KEYWORDS GSS
SOURCE thale cross.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Etrypophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis 1 (bases 1 to 1018)
REFERENCE Fennel, J., Dewar, K., Buehler, F., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.
AUTHORS BAC End Sequences at ATCC
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: F26F4-Sp6, F26F4-T7, F26F4-Sp6.1
Contact: Ecker, J.
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 288
High quality sequence stop: 355.
Location/Qualifiers
1..1018
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F26F4"
/clone_11b="1G"
/sex="hermaphrodite"
/note="Vector: ReliOnAC1; Site_1: EcoRI; Site_2: EcoRI; Produced by Thomas Altmann"
BASE COUNT 338 a 192 c 163 g 298 t 27 others
ORIGIN

Alignment Scores:
Pred. No.: 186 Length: 1018
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
Gaps: 0
DB: 17

US-09-831-458a-12 (1-325) x B12648 (1-1018)
Q7 211 SerClaraGASrPhlissAspserVal 219
|||||
Db 614 TCCTAAGAAACATGCTATCATCTGTT 640

RESULT 42
CNS00385 1101 bp DNA linear GSS 04-JUN-1999
DEFINITION BAC080624 of PPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION CNS00385
VERSION AL063897.1 GI:4941754
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Euryptera; Neoptera; Endopterygota; Diptera; Brachycera; Muscivora; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : 4p 191 41006 EVRY code FRANCE (E-mail: seq@genoscope.cns.fr - Web: www.genoscope.cns.fr)
COMMENT Determination of this BAC end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library was prepared by Kalliope Coscova and Aaron Mammossier in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named PPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BAC080624"
/clone_11b="PPCI-98"
/note="end : 47"
BASE COUNT 223 a 238 c 232 g 374 t 34 others
ORIGIN

Alignment Scores:
Pred. No.: 206 Length: 1101
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
Gaps: 0
DB: 17

US-09-831-458a-12 (1-425) x CNS00385 (1-1101)
Q7 272 ProGASrPhlissAspserVal 280
|||||
Db 191 CCACTTCAACCTCTCCAGGCTTAT 217

RESULT 43
AF326918 3108 bp mRNA linear HTG 13-JUL-2001
DEFINITION Homo sapiens metalloproteinase-disintegrin meltrin beta (FMS34)
ACCESSION AF326918
VERSION AF326918.1 GI:12276179
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Cladacea; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 3108)
AUTHORS Wang, Y.-G. and Gong, J.
TITLE Identification of FMS34, a novel human gene encoding for metalloproteinase-disintegrin meltrin beta unpublished
JOURNAL 2 (bases 1 to 3108)
Wang, Y.-G.
Direct Submission
Submitted (07-DEC-2000) Beijing Fengkesheng Function Gene Technology Ltd., 4 Tong Tiao Lu Chang Street, Xuanwu District, Beijing 100050, P.R. China
Location/Qualifiers
1..3108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
1..3108
gene

CDIS

/gene="FKSG34"
46...2916
/gene="FKSG34"
/codon_start=1
/product="metalloproteinase-disintegrin mectin beta"
/protein_id="AA050282.1"
/db_xref="GI:12276180"
/translation="MPGAGARLCLLAALPLPPRAKEPWTSGSESGKLOH
ELIPWKTSESPREKHPLEALVMEAGRELILDEKNEQLFAFSTETHTYSGN
PPTTRKLEDEHFGYGTRETELSSVTLSGTGILITSSNLSEVLEPDSGON
LYRSEHLKPPGNGGFEHSPPTTDMALQFTQTKRPPKREELNSMVELEYIV
ADLEFQKRRRQDQATKHLPLIANYVKEFESLNIPLALYCEVETGNNCEVSEN
YNTLSEFLSKRRKLLAQKYHDAQILITGMSFTGTTIGLAPLMAMCSVOSGVNDRS
ENALGVAAIHAHEMGHFGMTDSDADCCASADGCTMAATGHPPEKPVNGCNRE
LRLVLOSGGGKCLSNMPTDRLVGGRRQNGYLEDGEGDCGEECECNPCNANCT
LRPGACAGSGCHQCKLAPGLTLCREARAPDLPECTGKSPHCTNFMODGTPCE
GGAYCYNGMCLTYGOCQOLMGARAPAPDLCEKVVAGDTPGCKVNGEHRK
NMADAKCKLQCSSEARPLESNAPVITITIMNGROIQCGTHYRGPEEGMLNP
GIWMTGCKGVNHICFEGCGRTSFEETGCGKCKNGKNGVCNNONCICLPGMAPPC
NITPHGSGIDSGMPPESEYGVAVLVAIILVAILMYTCCKNNKLGQJLPSALP
SKROOFSCPEFVSQNSGTGHANPTFKLQTPGKRVINTPELLKPSQPPRPDPY
LRGSPAPLPALHLSAANSPGSGOIERFTSSPPSPPLPIPAEPCIVSDFSRP
RPPKALPANVPYGRSLPRPGASPLRPGAGPQSPPLAIAFKVSPRALKVKAG
TRCIGGRCVKEKTKOFLVYVTELPKPKAKHSCFLVPA"

BASE COUNT 716 a 886 c 895 g 611 t

ORIGIN

Alignment Scores:
Pred. No.: 799 Length: 3108
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
DB: 11 Gaps: 0

US-09-831-458a-12 (1-325) x AF326918 (1-3108)

QY 34 ALaGyValLeuValAlaLeuVal 42
|||||
Db 2161 GCTGACTGTGTGGCCATCTGTGIG 2187

RESULT 44
BM441797 164 bp mRNA linear EST 23-JUL 2002
LOCUS Ebed07_S0001_108_R endosperm, 28 DPA, no treatment, cv Optic,
DEFINITION Ebed07 Hordeum vulgare cDNA clone Ebed07_S0001_108 5', mRNA
sequence.
ACCESSION BM441797 GI:21932628
VERSION BM441797.2 GI:21932628
KEYWORDS EST.
SOURCE Hordeum vulgare.
ORGANISM Hordeum vulgare.
REFERENCE 1 Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardie, L.,
AUTHORS Ramsay, L., Machray, G., Marshall, D.F.M., and Waugh, R.,
TITLE Development of Barley Transcriptome Resources
JOURNAL Unpublished (2001)
COMMENT on Feb 1, 2002 this sequence version replaced q1:18472572.
Contact: Waugh, R., Marshall, D.F.
Genome Dynamics/Computational Biology
Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5PA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: estescr@sari.ac.uk
ALL sequence has a phred quality score of 20 or over
Seq primer: M13 reverse.
FEATURES
1. 164
Location/Qualifiers
/organism="Hordeum vulgare"

CDIS

/cultivar="optic"
/db_xref="taxon:4513"
/clone="Ebed07_S0001_108"
/clone_1lb="endosperm, 28 DPA, no treatment
Ebed07"
/tissue_type="endosperm"
/db_stage="28 DPA"
/lab_host="p1108"
/note="Vector: pSPMT1; Site 1: Sal I; Site 2: Xba I; Site 3: Xba I. Non-normalised library, directionally cloned from endosperm tissue dissected from 28 days post anthesis in glasshouse plants. Developed as part of the barley resources of HRSR/SPEKAP funded cereal Gene Function project."

BASE COUNT 55 a 31 c 45 g 43 t

ORIGIN

Alignment Scores:
Pred. No.: 188 Length: 164
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: 13 Gaps: 0

US-09-831-458a-12 (1-325) x BM441797 (1-164)

QY 23 LeuValLeuGlnLeuLeuSerPhe 40
|||||
Db 81 CTGATTATACACTCTGTGTCATT 58

RESULT 45
BQ371071 171 bp mRNA linear
LOCUS BQ371071
DEFINITION KC2-FN0990-060600-011-a05 FN0990 Homo sapiens cDNA
ACCESSION BQ371071.1 GI:21046585
VERSION BQ371071.1 GI:21046585
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida,
AUTHORS Nagai, M.A., da Silva, W., Jr., Zabo, M.A., Bordin, S.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jonsson,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza,
Simpson, A.J.,
TITLE Shotgun sequencing of the human transcriptome with
JOURNAL sequence tags
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 4491-4496 (2000)
COMMENT 2620263
Contact: Simpson, A.J.J.,
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Ria Prof Antonio Prudente, 109, 4 andar, 01506-91
Brazill
Tel: 55-11-2704922
Fax: 55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human
Project. This entry can be seen in the following URL:
(http://www.ludwig.org.br/cgi-bin/ftp/getntm2.pl?nt=pe
060600-011-a05&ts=2000-06-04-1)
Seq primer: puc 18 forward
High quality sequence starts: 9
High quality sequence stop: 171.
Location/Qualifiers
1. 171
/organism="Homo sapiens"
/db_xref="taxon:9606"

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 13:53:27 ; Search time 2755 seconds

(without alignments)
3433.180 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325
Sequence: 1 MSRSKPPWQQLGLGCLGH.....KQDVNWICKKPAACEFHE 325

Scoring table:
OLIGO
Xgapop 60.0 ; Xgapext 60.0
Ygapop 60.0 ; Ygapext 60.0
Kgapop 6.0 ; Kgapext 7.0
Delop 6.0 ; Deltext 7.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 5

Total number of hits satisfying chosen parameters: 581759

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL-frame-p2n.model -DEV=xlh
-O/cgnt2.1/USPTO.spool/US09831458/runat_05122002_103421_16305/36f-4acry.fasta.1.519
-DB-GenBml -QFMT-fastap -SUFFIX=oli.rge -MINMATCH=0 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=humand0.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=5 -ALIGN=45 -MODE=LOCAL -OUTFMT=pro
-NORM-ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09831458.acrg.1.1.1616 -runat_05122002_103421_16305 -NPRO=6 -TPRO 3
-NO_XLPEXT -NO_MMAPP -THRESDUERT -NEG_SCORES=0 -WAIT -LON=0.05 -DEV=1146000000
-MARKTIME0001-30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELop=6 -DELEXT=7

Database :

GenBml :
1: qb.ba:*
2: qb.htg:*
3: qb.in:*
4: qb.om:*
5: qb.ov:*
6: qb.pat:*
7: qb.ph:*
8: qb.pl:*
9: qb.pr:*
10: qb.ro:*
11: qb.sts:*
12: qb.sy:*
13: qb.un:*
14: qb.vl:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.mu:*
21: em.ov:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vl:*
30: em.htg.hum:*
31: em.htg.in:*
32: em.htg.other:*
33: em.htg.mus:*
34: em.htg.ph:*
35: em.htg.ro:*
36: em.htg.mam:*
37: em.htg.vl:*
38: em.sy:*
39: em.htg.hum:*
40: em.htg.mus:*
41: em.htg.other:*

Pred. No. is the number of results predicted by chance 1
score greater than or equal to the score of the result
and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Accession	Score
1	168	51.7	1643	6	AX287118	AX287118
2	168	51.7	1643	9	AB015629	AB015629
3	166	51.1	1200	9	AY042244	AY042244
4	166	51.1	1800	9	AF290867	AF290867
5	153	47.1	1510	9	AF245219	AF245219
6	153	47.1	1535	9	AY042236	AY042236
7	153	47.1	1648	9	AY042245	AY042245
8	138	42.5	792	9	AY042240	AY042240
9	138	42.5	999	9	AY042238	AY042238
10	125	38.5	1472	9	AY042239	AY042239
11	125	38.5	4506	9	AF209480S2	AF209480S2
12	125	38.5	143619	9	AC008812	AC008812
13	120	36.9	811	9	AY042237	AY042237
14	115	35.4	190628	2	AC024700	AC024700
15	97	29.8	2991	9	AK024750	AK024750
16	78	24.0	347	9	AY042233	AY042233
17	78	24.0	939	9	AY042234	AY042234
18	78	24.0	994	9	AY042232	AY042232
19	78	24.0	1083	9	AY042236	AY042236
20	78	24.0	1105	9	AY042231	AY042231
21	78	24.0	1143	9	AY042235	AY042235
22	78	24.0	1197	9	AY042232	AY042232
23	78	24.0	1212	6	AR166109	AR166109
24	78	24.0	1212	6	AX42004	AX42004
25	78	24.0	1215	9	AX049404	AX049404
26	78	24.0	1215	9	AY042231	AY042231
27	78	24.0	1243	9	AX042230	AX042230
28	78	24.0	1312	6	AX287104	AX287104
29	78	24.0	1312	9	HDMLCCT1	HDMLCCT1
30	78	24.0	1315	9	AY042229	AY042229
31	78	24.0	4266	9	AF290866	AF290866
32	78	24.0	4424	9	AF209479	AF209479
33	78	24.0	143619	9	AC008812	AC008812
34	78	24.0	177962	9	AC008764	AC008764
35	69	21.2	807	9	AY042227	AY042227
36	43	13.2	1212	6	AR166106	AR166106
37	43	13.2	1212	6	AX342064	AX342064
38	43	13.2	1212	9	AF091085	AF091085
39	40	12.3	1146	9	AF090755	AF090755
40	40	12.3	1166	9	AY040319	AY040319
41	35	10.8	1212	6	AR166107	AR166107
42	35	10.8	1212	6	AX42005	AX42005
43	34	10.5	1146	9	AF343727	AF343727
44	26	8.0	1215	9	AF091086	AF091086
45	21	6.5	507	9	AY042224	AY042224

RESULT 1

ALIGNMENTS

AX:87118
 LOCUS AX287118 1643 bp DNA linear PAT 21 NOV 2001
 DEFINITION Sequence 16 from Patent WO0164752.
 ACCESSION AX287118
 VERSION AX287118.1 GI:17049092
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homo; Homininae; Homo sapiens.
 REFERENCE
 1. Littman, D. P., Kwon, D., van Kester, Y. Q., and Goldfarb, R. H. Antibody Inhibiting the Binding between gp120 and Receptor and Screening Methods Patent: WO 0164752-A (1997-SEP-1901).
 JOURNAL NEW YORK UNIVERSITY (US); KATHOLIEKE UNIVERSITEIT NIMEGEN (NL)
 FEATURES
 source
 1. 1643
 Location/Qualifiers
 CDS
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 25..915
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD12298.1"
 /db_xref="GI:17049093"
 /translation="MSDSKEPRVQGLICGAGLVLLSLMLAGVVALVWY
 SKVPSLSQESQDAIYONLTOKAAGVLSKSKLOEIYQELTLKAAGVLPERS
 KLOEIYQELTLKAAGVLPERSKLOEIYQELTLKAAGVLPERSKLOEIYQELTL
 KAAGVLPERSKLOEIYQELTLKAAGVLPERSKLOEIYQELTLKAAGVLPERS
 QOOIYQELTLKTAERLCRCRCPKDWTFPQNCYFMSNSQNMHDSVIAQEVRAQIV
 VKTAEEDLPVILQWRTQO"
 BASE COUNT 388 a 429 c 440 g 386 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,446-166 Length: 1643
 Score: 168.00 Matches: 168
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 51.69% Indels: 0
 DB: 6 Gaps: 0
 US-09-831-458a-12 (1-325) x AX287118 (1-1643)
 QY 1 MetSerAspSerIyscIuPrArqVtGIncInleuGlyLeuLeuGlycysIeuGlyHis 20
 Db 25 ATGAGTACCTCCAAAGCAACCAAGAGTGTAGTACCTGGGCTCTGGGCTGTGGCTAT 84
 QY 21 GYALALEUVALLEUENLEUCUSERPHEMETLEUENLAAGLYVALLEUVALALE 40
 Db 85 GGGGCGCTGGTGTGCACT 144
 QY 41 LeuValGlnValSerIysValProSerSerLeuSerGlnGlnInSerGlnGlnAspAla 60
 Db 145 CTTCCTCAACTGTCTCAAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
 QY 61 LLeTyrGlnAsnLeuThrGlnLeuLeuLysAlaAlaValGlyGlnLeuSerGlnLysSerIys 80
 Db 205 ATTATTAATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 264
 QY 81 LeuGlnGlnGlnLeuTyrGlnGlnGlnLeuThrGlnLeuLeuLysAlaAlaValGlyGln 100
 Db 265 GGGGCGCTGGTGTGCACT 144
 QY 101 LysSerIysLeuGlnGlnGlnLeuTyrGlnGlnGlnLeuThrGlnLeuLeuLysAlaAlaValGlyGln 120
 Db 325 AATTCCTCAACTGTCTCAAGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 384
 QY 121 LeuProGlnLysSerIysLeuGlnGlnGlnLeuThrGlnLeuLeuLysAlaAlaValGlyGln 140
 Db 385 TTCTATTAATTAATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 444
 QY 141 ValGlyValLeuProGlnLysSerIysLeuGlnGlnGlnLeuThrGlnLeuLeuLysAlaAlaValGlyGln 160

Db 445 GGGGCGCTGGTGTGCACT 504
 QY 161 LysAlaAlaValGlnGlnGlnLeuPro 168
 Db 505 AAGCTGTCAAGTGTGAGTGTGCACT 528
 RESULT 2
 AB015629 1643 bp mRNA linear PAT 26 JUL 2001
 LOCUS
 DEFINITION Homo sapiens mRNA for type II membrane protein similar to HIV
 gp120 binding C-type lectin, complete cds. clone: HPO1347.
 ACCESSION AB015629
 VERSION AB015629.1 GI:4586835
 KEYWORDS type II membrane protein similar to HIV gp120-binding C-type
 lectin.
 SOURCE Homo sapiens liver cDNA to mRNA, clone: HPO1347.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE
 1. Yokoyama-Kobayashi, M., Yamaguchi, T., Setiye, S., and Kato, S. Selection of cDNAs encoding putative type II membrane proteins on the cell surface from a human full-length cDNA bank
 Gene 228 (1-2), 151-162 (1999)
 JOURNAL MEDLINE
 99173880
 2 (bases 1 to 1643)
 AUTHORS Kato, S.
 TITLE Direct Submission
 JOURNAL Submitted (16-JUN-1998) Seishi Kato, Research Institute of National
 Rehabilitation Center for the Disabled, Department of
 Rehabilitation Engineering, 4-1 Namiki, Tokorozawa, Saitama
 359-8555, Japan (E-mail: seishi@rehab.go.jp,
 Tel: 042-995-3100 (ex. 2568), Fax: 042-995-3152)
 FEATURES
 source
 1. 1643
 Location/Qualifiers
 CDS
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HPO1347"
 /translation="MSDSKEPRVQGLICGAGLVLLSLMLAGVVALVWY
 SKVPSLSQESQDAIYONLTOKAAGVLSKSKLOEIYQELTLKAAGVLPERS
 KLOEIYQELTLKAAGVLPERSKLOEIYQELTLKAAGVLPERSKLOEIYQELTL
 KAAGVLPERSKLOEIYQELTLKAAGVLPERSKLOEIYQELTLKAAGVLPERS
 QOOIYQELTLKTAERLCRCRCPKDWTFPQNCYFMSNSQNMHDSVIAQEVRAQIV
 VKTAEEDLPVILQWRTQO"
 BASE COUNT 388 a 429 c 440 g 386 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,446-166 Length: 1643
 Score: 168.00 Matches: 168
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 51.69% Indels: 0
 DB: 6 Gaps: 0
 US-09-831-458a-12 (1-325) x AB015629 (1-1643)
 QY 1 MetSerAspSerIyscIuPrArqVtGIncInleuGlyLeuLeuGlycysIeuGlyHis 20
 Db 25 ATGAGTACCTCCAAAGCAACCAAGAGTGTAGTACCTGGGCTCTGGGCTGTGGCTAT 84
 QY 21 GYALALEUVALLEUENLEUCUSERPHEMETLEUENLAAGLYVALLEUVALALE 40
 Db 85 GGGGCGCTGGTGTGCACT 144

Db 85 GGGCGCTGGTGGCAACTCTCTCTTCATGCTTGGCTGGGGTCCCTGGCGCCATC 144

QY 41 LeuValGlnValSerLysValProSerSerLeuSerGlnGlnGlnSerGlnGlnAspAla 60

Db 145 CTGTGCGAAGTGTCCCAAGTCCCACTCCCTCAAGTCAGAGAACATCCGAGCAATGCA 204

QY 61 TLeuGlnAsnLeuThrGlnLeuLysAlaAlaValGlyLeuSerGlnLysSerLys 80

Db 205 ATTAACCTAGAACCTGACCTGACCTTAAACCTGACCTGACCTGACCTGACCTGACCT 264

QY 81 LeuGlnGlnLeuLeuGlnGlnLeuThrGlnLeuLysAlaAlaValGlyLeuProGln 100

Db 265 CTGACAGAGATTCACAGAGATTCACAGAGATTCACAGAGATTCACAGAGATTCACAG 324

QY 101 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120

Db 325 AAATGCAAGCTGCAAGAT 384

QY 121 LeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 140

Db 485 TTGTCAGAGAAATCCAGAGATTCACAGAGATTCACAGAGATTCACAGAGATTCACAG 444

QY 141 ValGlyGlnLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 160

Db 445 GTGGCTGAGTGGCCAGAGAAATCCAGAGATTCACAGAGATTCACAGAGATTCACAG 504

QY 161 LysAlaAlaValGlyLeuPro 168

Db 505 AAGCTTCTATGTCCTGATTTGCA 528

RESULT 3
AY042234 1200 bp mRNA linear FRI 30 AUG 2001
LOCUS AY042234
DEFINITION Homo sapiens MDC-STGN2 type I isoform (CD209L) mRNA, complete cds,
alternatively spliced.
ACCESSION AY042234
VERSION AY042234.1 GI:15383695
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Cladista; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homini; Homo.
REFERENCE 1 (bases 1 to 1200)
Mummidu, S., Catalano, G., Lam, L., Hoefle, A., Telles, V., Hegun, K.,
Jimenez, J., Ahuja, S. S. and Ahuja, S. K.
Extensive repertoire of membrane-bound and soluble dendritic
cell-specific ICAM-3-grabbing nonintegrin 1 (DC-STGN1) and DC-STGN2
isoforms: inter-individual variation in expression of DC-STGN
transcripts
JOURNAL J. Biol. Chem. 276 (45), 33195-33212 (2001)
MEDLINE 2141847
PubMed 11337487
REFERENCE 2 (bases 1 to 1200)
Mummidu, S. and Ahuja, S. K.
Direct Submission
Submitted (03-JUN-2001) Medicine, South Texas Veterans Health Care
System and University of Texas Health Science Center at San
Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229, USA
FEATURES
SOURCE
1..1200
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="19p13.3"
1..1200
/gene="CD209L"
1..1200
/gene="CD209L"
/gene="CD209L"
/note="similar to the nucleotide sequence deposited in
GenBank Accession Number AC008812; alternatively spliced"
/product="MDC-STGN2 type I isoform"
/protein_id="AAK91859.1"

/db_xref="GI:15383695"
/translation="MDSKRPVVOUGLLEEDPTSGRI...
CISGALVLDLSTLAVLAVLAVLVGVSVSSLSQSPF...
LSEKSKLOEIVQELTOLKAVALPEPSKLOEIVQELT...
ELTRKAAY-PLPEPSKLOEIVQELTOLKAVALPEPSK...
PEPSKLOEIVQELTOLKAVALPEPSKLOEIVQELT...
GNCYFMSNSORHMDSVTACQVRAQIVLKTAEKQNF...
QERTWQVWVPSPI SPSPFVRWNSCFPNNNSNCFAP...
AACRFDE"

BASE COUNT 317 a 313 c 44 g 226 t

ORIGIN

Alignment Scores:

Pred. No.:	4,42e-164	Length:	1200
Score:	166.00	Matches:	23*
Percent Similarity:	98.31%	Conservative:	23*
Best Local Similarity:	98.31%	Mismatches:	2
Query Match:	51.08%	Indels:	4
DB:	9	Gaps:	0

05-09-831-458a-12 (1 325) x AY042234 (1 1200)

QY 91 LeuLysAlaAlaValGlyLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGln 120

Db 493 CTGAAGCTGCGAGTGGAGATTCACAGAGATTCACAGAGATTCACAGAGATTCACAG 553

QY 111 LeuThrLeuLysAlaAlaValGlyLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGln 120

Db 553 CTGAAGCTGCGAGTGGAGATTCACAGAGATTCACAGAGATTCACAGAGATTCACAG 613

QY 131 TyrGlnGlnLeuThrArg-LeuLysAlaAlaValGlyLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGln 120

Db 613 TACCAAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAG 673

QY 150 GlnGlnGlnLeuThrArg-LeuLysAlaAlaValGlyLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGln 120

Db 673 CTGAAGCTGCGAGTGGAGATTCACAGAGATTCACAGAGATTCACAGAGATTCACAG 731

QY 170 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120

Db 731 ACTCCAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAG 791

QY 190 GlnGlnGlnLeuThrArg-LeuLysAlaAlaValGlyLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGln 120

Db 791 TACCAAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAG 851

QY 210 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120

Db 851 ACTCCAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAG 911

QY 230 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120

Db 911 TACCAAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAG 971

QY 250 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120

Db 971 TACCAAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAG 1031

QY 270 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120

Db 1031 TACCAAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAG 1091

QY 290 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120

Db 1091 TACCAAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAG 1151

QY 310 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 120

Db 1151 TACCAAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAGATTCACAG 1211

RESULT 4
AF290887 1200 bp mRNA linear
LOCUS AF290887

DEFINITION	Homo sapiens L-Ston mpna, complete cds
ACCESSION	AJ290887
VERSION	AJ290887.1 GI:13383469
KEYWORDS	.
SOURCE	Homo sapiens
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
ADDITIONAL	1 (bases 1 to 1800) Bashirova,A.A., Geijtenbeek,T.B.H., van Duinhoven,G.C.F., van Vlier,S.L., Filtinger,J.R.G., Martin,M.P., Wu,L., Martin,D.P., Viebig,N., Knolle,P.A., KewalRamani,V.N., Van Kooyk,Y. and Farrington,M. A dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin (DC-STON)-related protein is highly expressed on human liver sinusoidal endothelial cells and promotes HIV-1 infection J. Exp. Med. 193 (6), 671-678 (2001)
JOURNAL	J. Exp. Med. 193 (6), 671-678 (2001)
MEDLINE	21157496
PUBMED	11257134
REFERENCES	2 (bases 1 to 1800)
AUTHORS	Bashirova,A.A.
TITLE	Direct Submission
JOURNAL	Submitted (27-JUL-2000) LRP, NCIFRFP, P.O. Box B, Frederick, MD 21702, USA
FEATURES	Location/Qualifiers
SOURCE	1..1800 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="1p" /map="1p13.3-p13.2" 54..1184 /note="type II membrane C-type lectin"
CDS	/product="L-Ston" /protein_id="AAK20938.1" /db_xref="GI:13383470" /translation="MSDKSEPPVOGLLEEDPTSGILPEPPOFOPIHQBSSTY CLGHAGLVOLIEMFLAGYLVALGVSVSPSSLDQSDPDATLYNLTOLKAANGG LSKKSLDLOIYOLTOLKAANGRIPEKSKIOPLVGHITAKAVDELIPKSKIETPG EHLRLAKAANGRIPEKSKIQPYGTETIRLKAVNELPKSKIKGEYGPRTTLKAAVEI PHOSKOOLYOEELDLKTAEYRLRHCPKWTEFGNGTFFKNSSPNHNHSVTVAAPV PAVLVNTIAEDNFLQLQTSSNPFSMWGLSELNLNEGITWAWVSPIPSSEDEYWS GEENSGNDIAEFNSGNVNRCDVDNVICKKPACFRDE"
RANGE COUNT ORIGIN	435 a 475 c 468 g 422 t
Alignment Scores:	
Prod No :	5 92e-164 Length: 1800
Score:	166.00 Matrices: 233
Percent Similarity:	98.31% Conservative: 0
Best Local Similarity:	98.31% Mismatch: 2
Query Match:	51.08% Indels: 4
DIM:	9 Gaps: 0
US-09-831-458A-12 (1-325) x AF290887 (1-1800)	
UY	100 ysgalslaalaavaglglyglnleuprcgluysserylskeaglmgultvtyingln 110
DB	CTGAAGGCTTAATGGATTAGTTGATAVAAAAGTGTCAGCAGATCAACAAGA 636
UY	111 leathlvrlrldysAlaaAvatgcylglylrllrrlrrtutulyssstgslawdlmalle 160
DB	CTGACCCTCATGAAGGCTGCACACTTCACTTAGTCCAGGAATCCAAGTCCAGGAGATC 596
UY	131 tyfqlndltdeuthafay-leouygsAlaalavaglglyluemprccglutyssotrgsla 150
DB	TACCGAGAGCTGACGA-GGTGAAGAGCTGCATGTGGTAGTG-CAGAGAAATCCAAAGT 655
UY	150 uahncvtclctgytlacgtatufhafrq-leouygsAlaalavaglglyluemprccspasps 170
DB	GAAATGATATACCAAGAGCTAACCA-GCTGAAGAGCTGTGATGATGATGATGATGATGATG 714

QY	170	QY	175	QY	178	QY	179	QY	180	QY	181	QY	182	QY	183	QY	184	QY	185	QY	186	QY	187	QY	188	QY	189	QY	190	QY	191	QY	192	QY	193	QY	194	QY	195	QY	196	QY	197	QY	198	QY	199	QY	200	QY	201	QY	202	QY	203	QY	204	QY	205	QY	206	QY	207	QY	208	QY	209	QY	210	QY	211	QY	212	QY	213	QY	214	QY	215	QY	216	QY	217	QY	218	QY	219	QY	220	QY	221	QY	222	QY	223	QY	224	QY	225	QY	226	QY	227	QY	228	QY	229	QY	230	QY	231	QY	232	QY	233	QY	234	QY	235	QY	236	QY	237	QY	238	QY	239	QY	240	QY	241	QY	242	QY	243	QY	244	QY	245	QY	246	QY	247	QY	248	QY	249	QY	250	QY	251	QY	252	QY	253	QY	254	QY	255	QY	256	QY	257	QY	258	QY	259	QY	260	QY	261	QY	262	QY	263	QY	264	QY	265	QY	266	QY	267	QY	268	QY	269	QY	270	QY	271	QY	272	QY	273	QY	274	QY	275	QY	276	QY	277	QY	278	QY	279	QY	280	QY	281	QY	282	QY	283	QY	284	QY	285	QY	286	QY	287	QY	288	QY	289	QY	290	QY	291	QY	292	QY	293	QY	294	QY	295	QY	296	QY	297	QY	298	QY	299	QY	300	QY	301	QY	302	QY	303	QY	304	QY	305	QY	306	QY	307	QY	308	QY	309	QY	310	QY	311	QY	312	QY	313	QY	314	QY	315	QY	316	QY	317	QY	318	QY	319	QY	320	QY	321	QY	322	QY	323	QY	324	QY	325	QY	326	QY	327	QY	328	QY	329	QY	330	QY	331	QY	332	QY	333	QY	334	QY	335	QY	336	QY	337	QY	338	QY	339	QY	340	QY	341	QY	342	QY	343	QY	344	QY	345	QY	346	QY	347	QY	348	QY	349	QY	350	QY	351	QY	352	QY	353	QY	354	QY	355	QY	356	QY	357	QY	358	QY	359	QY	360	QY	361	QY	362	QY	363	QY	364	QY	365	QY	366	QY	367	QY	368	QY	369	QY	370	QY	371	QY	372	QY	373	QY	374	QY	375	QY	376	QY	377	QY	378	QY	379	QY	380	QY	381	QY	382	QY	383	QY	384	QY	385	QY	386	QY	387	QY	388	QY	389	QY	390	QY	391	QY	392	QY	393	QY	394	QY	395	QY	396	QY	397	QY	398	QY	399	QY	400	QY	401	QY	402	QY	403	QY	404	QY	405	QY	406	QY	407	QY	408	QY	409	QY	410	QY	411	QY	412	QY	413	QY	414	QY	415	QY
----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----

Q7	144	LeptodactylusSteirysLeucingitullyPringituleantPravglouySAliaVid	163
D9	69048	TTGSCCAAGAAATCTAATCTTGTAGACAGATTTCATCAGGAAGTGGACCTGCGACTGCA	69107
Q7	164	ValechyluleuPro	168
D9	69108	GTAATATCATGATCCCA	69122
RESULT 13			
AD042237			
DEFINITION	AY042237 Homo sapiens CDC-STGN2 type VI isoform (CD209L) mRNA, complete cds.	811 bp	mRNA linear pri 30-AUG-2001
ACCESSION	AY042237		
VERSION	AY042237		
KEYWORDS	AY042237.1 GI:15383611		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 811)		
JOURNAL	Mummidl,S., Catarao,G., Lam,L., Hoefle,A., Telles,V., Begum,K.,		
PUBLISHED	Jimenez,F., Ahuja,S.S. and Ahuja,S.K		
EXTENSIVE REPERTOIRE OF MEMBRANE-BOUND AND SOLUBLE DENDRITIC CELL-SPECIFIC CDAM-1 GRABBIN NOINTERGRIN 1 (DC-STGN1) AND DC-STGN2 ISOFORMS, INTER INDIVIDUAL VARIATION IN EXPRESSION OF DC-STGN TRANSCRIPTS			
J. Biol Chem.	276 (45), 33196-33212 (2001)		
MEDLINE	11337487		
PUBMED	11337487		
REFERENCE	2 (bases 1 to 811)		
AUTHORS	Mummidl,S. and Ahuja,S.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (03-JUN-2001) Medicine, South Texas Veterans Health Care System and University of Texas Health Science Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229, USA		
FEATURES	Location/Qualifiers		
source	1..811		
gene	/organism="Homo sapiens"		
CDS	/db_xref="taxon:9606"		
	/chromosome="19"		
	/map_ "19p13.3"		
	1..811		
	/feature="cd209L"		
	1..699		
	/gene="CD209L"		
	/note="similar to the nucleotide sequence deposited in GenBank Accession Number AC009812; alternatively spliced"		
	/codon_start=1		
	/product="cdc-stgn2 type VI isoform"		
	/protein_id="AAK91862.1"		
	/db_xref="GI:15383612"		
	/translation="MSDSKPRPVQLGLEEDPTTSIRLPFDGFQDIHQSKSSSTGLGHGAVIOLSLFMLAGVAIVAIIVGVSPVSLSDSGEQUAIYONILVIAKAAGVLSKRSKLOEIYOELTOLKAANGELPEKSLOEIYOETRIKAAVETIPKSKIIEIYVLTOKLAVERFCPRCPKIWTFFPGSNCFPSNSCPKMHSVTACQCVVAGIVITATAEOLPAVEQWKIQP"		
BASE COUNT	213 a 218 c 226 g 154 t		
ORIGIN			
Alignment Scores:			
Prod. No.:	5.5e-116	Length:	811
Score:	120.00	Matches:	120
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	36.92%	Indels:	0
DB:	9	Gaps	0
OS:09-831-458A-12 (1-325) x AY042237 (1-811)			
Q7	16	GTPYSLEUHLISCLYALALCALVALLEUCINLEUESEPHNCTEULEVALACTY	35

D6	140	GAGGTCTTTCCTCAATACGAGCGCCTGGTGTTGAACACTCCCTCTTCAATCATCTCTTAACTGAGG	189
O7	36	VallentValAlalleleVallelchValSerLysValProSerSerLeuSerGlnGluIn	55
D6	190	GGGCGGCTGGAGGACCTGGGTAAGAGGCTAACAGGCTAAGCTTCTTAAGTAAAGAGCAA	249
OY	56	SerGlnGlnAsnAlaAlaIleTyrlChlaSuleqThGlnleuLSAlaAlaValIslgIleu	75
D6	250	TGGTAGCAAGAGCGAAMGATACCAACAACCTGAGCTAAGCTTAAGCTTCAAGTCAAGCTC	309
OY	76	SerLelLysSerLysLeuGlrectuIlleTyrlrlnGlnleuThrGlnleuLysAlaValal	95
D6	310	TGAAATAAATATTAAGTGGTAAGAGATACAGAGAGCTGAGCTTAAAGTAAAGTAAAGG	369
OY	96	GIAG	115
D6	470	GGTGGATTCGTAACAAGAATATTAACCTGGAGAGAGATGAGAGAGAGAGAGAGAGAGAG	429
O7	116	AlAlAValcIgcLc	135
D6	440	GGTGGATTCGTAACAAGAATATTAACCTGGAGAGAGATGAGAGAGAGAGAGAGAGAGAG	489

RESULT 14
AC024790/c

Locus	AC024790	190628 bp	UNA	linear	HUG 07-JUL-2000
DEFINITION	Homo sapiens chromosome 19 clone RP1142J18, WORKING DRAFT				
ACCESSION	AC024790				
VERSION	AC024790.4	CL8269248			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens.				
REFERENCE	Bakayev, M.I., 2000, Chordata, Crustacea, Vertebrata: Buteleostomi; Mammalia: Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 190628)				
TITLE	Waterston,R.H.				
JOURNAL	The sequence of Homo sapiens clone				
REFERENCE	2 (bases 1 to 190628)				
AUTHORS	Waterston,R.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-MAR-2000) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA				
COMMENT	On Jan 15, 2003 this sequence version replaced q1.7523915.				

Center: Washington University Genome Sequencing Center
 Contact code: WUGSC
 Web site:http://genome.wustl.edu/gsc/index.shtml

Project Information
 Center project name: LNH0042J18

Summary Statistics
 Sequencing vector: M13; 100k
 Sequencing vector: plasmid; 04
 Chemistry: Dye-primer FT; 100k of reads
 Chemistry: Dye terminator BigDye; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 173667 bases at least Q40
 Consensus quality: 178661 bases at least Q30
 Consensus quality: 181172 bases at least Q20
 Insert size: 175000; agarose-tip
 Insert size: 187528; sum-of-contigs
 Quality coverage: 5.24 in Q20 bases; agarose-tip
 Quality coverage: 5.00 in Q20 bases; sum-of-contigs

* Note: This is a working draft sequence. It currently consists of 32 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will

```

* be preserved.
1 1244: contig of 1244 bp in length
1245 1344: gap of unknown length
1345 2396: contig of 1052 bp in length
2397 2497: gap of unknown length
2497 3585: contig of 1089 bp in length
3586 5158: gap of unknown length
5159 5259: contig of 1474 bp in length
5259 6624: gap of unknown length
6624 6725: contig of 1366 bp in length
6725 8402: gap of unknown length
8402 10172: contig of 1677 bp in length
10172 10273: contig of 1671 bp in length
10273 12510: gap of unknown length
12510 12611: contig of 2238 bp in length
12611 14735: gap of unknown length
14735 14835: contig of 2125 bp in length
14835 16443: gap of unknown length
16443 16543: contig of 1608 bp in length
16543 19269: gap of unknown length
19269 19370: contig of 2726 bp in length
19370 21690: gap of unknown length
21690 21791: contig of 2321 bp in length
21791 25543: gap of unknown length
25543 25644: contig of 3753 bp in length
25644 28329: gap of unknown length
28329 31639: gap of unknown length
31639 31740: contig of 3311 bp in length
31740 36567: gap of unknown length
36567 39593: contig of 4727 bp in length
39593 39694: gap of unknown length
39694 46199: gap of unknown length
46199 50494: contig of 3027 bp in length
50494 50595: gap of unknown length
50595 56367: gap of unknown length
56367 61903: contig of 5773 bp in length
61903 62004: gap of unknown length
62004 67159: gap of unknown length
67159 74272: contig of 5156 bp in length
74272 80579: gap of unknown length
80579 80680: contig of 6308 bp in length
80680 88540: gap of unknown length
88540 88641: contig of 7861 bp in length
88641 95518: gap of unknown length
95518 102484: contig of 6878 bp in length
102484 102585: gap of unknown length
102585 109469: contig of 6866 bp in length
109469 130222: gap of unknown length
130222 130323: contig of 6844 bp in length
130323 154890: gap of unknown length
154890 171709: contig of 2654 bp in length
171709 171810: gap of unknown length
171810 190628: contig of 16720 bp in length
190628 190628: contig of 18819 bp in length.

FEATURES
source
1. 190628
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="RP11-42J18"
/misc_feature
1. 1244
/note="assembly_name:Contig15"
1345. 2396
Location/Qualifiers
1. 190628
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19"
/clone="RP11-42J18"
/misc_feature
1. 1244
/note="assembly_name:Contig15"
1345. 2396

```

```

/misc_feature
2497. 3585
/note="assembly_name:Contig16"
/misc_feature
3586. 5158
/note="assembly_name:Contig17"
/misc_feature
5259. 6624
/note="assembly_name:Contig18"
/misc_feature
6725. 8401
/note="assembly_name:Contig19"
/misc_feature
8502. 10172
/note="assembly_name:Contig20"
/misc_feature
10273. 12510
/note="assembly_name:Contig21"
/misc_feature
12611. 14735
/note="assembly_name:Contig22"
/misc_feature
16544. 19269
/note="assembly_name:Contig23"
/misc_feature
19370. 21690
/note="assembly_name:Contig24"
/misc_feature
21791. 25543
/note="assembly_name:Contig25"
/misc_feature
25644. 28329
/note="assembly_name:Contig26"
/misc_feature
31740. 36567
/note="assembly_name:Contig27"
/misc_feature
39694. 46199
/note="assembly_name:Contig28"
/misc_feature
50494. 50595
/note="assembly_name:Contig29"
/misc_feature
56367. 61903
/note="assembly_name:Contig30"
/misc_feature
62004. 67159
/note="assembly_name:Contig31"
/misc_feature
74272. 80579
/note="assembly_name:Contig32"
/misc_feature
80680. 88540
/note="assembly_name:Contig33"
/misc_feature
88641. 95518
/note="assembly_name:Contig34"
/misc_feature
102585. 109469
/note="assembly_name:Contig35"
/misc_feature
130222. 130323
/note="assembly_name:Contig36"
/misc_feature
154890. 171709
/note="assembly_name:Contig37"
/misc_feature
171810. 190628
/note="assembly_name:Contig38"

```

```

BASE COUNT 4336 a 49517 c 49715 g 44333 t 3127 other
ORIGIN
Alignment Scores:
Pred. No.: 4,93e-109 Length: 190628
Score: 115.00 Matches: 115
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0

```

Query March: 35,388 Indexes: 0
 Ds: 2 Gaps: 0
 us-09-831-458a-12 (1-325) x AK023750 (1-2991)
 QY 44 ValSerLysValProSerSerLeuSerGlnGlnInSerGlnInAspAlaIleTyrcIn 63
 |||||
 Db 105942 GTCAGACATCTACGAGAGCTGACCGGCTGAGAGGCTGAGTGGGAGAGTGGCAGAG 105943
 |||||
 QY 64 AsnLeuThrGlnLeuLysAlaAlaValGlyIleLeuSerGlnLysSerLysLeuGln 83
 |||||
 Db 105932 AACATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 105933
 |||||
 QY 84 IleTyrcGlnGlnLeuThrGlnLeuLysAlaAlaValGlyIleLeuProGlnLysSerLys 103
 |||||
 Db 105872 ATCTACCAAGAGTGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 105873
 |||||
 QY 104 LeuGlnGlnIleTyrcGlnLeuThrGlnLeuLysAlaAlaValGlyIleLeuProGln 123
 |||||
 Db 105812 CTCAGACATCTACGAGAGCTGACCGGCTGAGAGGCTGAGTGGGAGAGTGGCAGAG 105813
 |||||
 QY 124 LysSerLysLeuGlnGlnIleTyrcGlnLeuThrGlnLeuLysAlaAlaValGlyIle 143
 |||||
 Db 105752 AAATCCAGCTGACGAGATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 105753
 |||||
 QY 144 LeuProGlnLysSerLysLeuGlnGlnIleTyrcGlnLeuThr 158
 |||||
 Db 105692 TTCTTAAAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 105693
 |||||
 RESULT 15
 AK023750 2991 bp mRNA linear FRI 01 AUG 2002
 Locus Homo sapiens cDNA FL113688 fls, clone PLACE2000062, highly similar
 DEFINITION
 Homo sapiens mRNA for type II membrane protein.
 ACCESSION
 AK023750.1 GI:10435777
 VERSION
 oligo capturing fls (full insert sequence)
 KEYWORDS
 Homo sapiens placenta cDNA to mRNA, clone_11b-PLACE2
 SOURCE
 clone: PLACE2000062
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
 Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
 Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,P., Takeuchi,K.,
 Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,T.,
 Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuno,Y. and Oshima,A.
 NEDO human cDNA sequencing project
 TITLE
 Unpublished
 JOURNAL
 2 (bases 1 to 2991)
 PREPARED
 Isogai,T. and Otsuki,T.
 AUTHORS
 Direct Submission
 JOURNAL
 Submitted (23-AUG-2000) Tokyo Isogai, Helix Research Institute,
 Genomics Laboratory, 152-3 Yoda Katsuta, Chiba 270-0312, Japan
 (E-mail:genomics@hri.co.jp, Tel:81-48-52-4975, Fax:81-48-52-3986)
 COMMENT
 NEDO human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan, cDNA full insert
 sequencing, Research Association for Biotechnology, cDNA library
 construction, 5'- & 3'-end one pass sequencing and clone selection;
 Helix Research Institute (supported by Japan Key Technology Center
 etc.) and Department of Virology, Institute of Medical Science,
 University of Tokyo.
 FEATURES
 SOURCE
 Location/Qualifiers
 1..2991
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="PLACE2000062"
 /tissue_type="Placenta"
 /tissue_11b="PLACE2"
 /note="cloning vector: pME18SFU3"
 8..658
 /note="unnamed protein product"

Alignment Scores:
 Pred. No.: 1,810-91 Length: 2991
 Score: 97.00 Matches: 124
 Percent Similarity: 98.41% Gaps: 0
 Best Local Similarity: 98.41% Mismatches: 1
 Query March: 29,858 Ids: 2
 Ds: 9 Gaps: 0
 us-09-831-458a-12 (1-325) x AK023750 (1-2991)
 QY 44 ValSerLysValProSerSerLeuSerGlnGlnInSerGlnInAspAlaIleTyrcIn 63
 |||||
 Db 53 GTCAGACATCTACGAGAGCTGACCGGCTGAGAGGCTGAGTGGGAGAGTGGCAGAG 112
 |||||
 QY 64 AsnLeuThrGlnLeuLysAlaAlaValGlyIleLeuSerGlnLysSerLysLeuGln 83
 |||||
 Db 113 AACCTGACCATCTTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 171
 |||||
 QY 83 IleTyrcGlnGlnLeuThrGlnLeuLysAlaAlaValGlyIleLeuProGlnLysSerLys 103
 |||||
 Db 172 GATCTTAAAGTGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 231
 |||||
 QY 104 LeuGlnGlnIleTyrcGlnLeuThrGlnLeuLysAlaAlaValGlyIleLeuProGln 123
 |||||
 Db 232 GTCAGACATCTACGAGAGCTGACCGGCTGAGAGGCTGAGTGGGAGAGTGGCAGAG 291
 |||||
 QY 124 LysSerLysLeuGlnGlnIleTyrcGlnLeuThrGlnLeuLysAlaAlaValGlyIle 143
 |||||
 Db 232 GTCAGACATCTACGAGAGCTGACCGGCTGAGAGGCTGAGTGGGAGAGTGGCAGAG 351
 |||||
 QY 143 LeuProGlnLysSerLysLeuGlnGlnIleTyrcGlnLeuThr 158
 |||||
 Db 352 GTCAGACATCTACGAGAGCTGACCGGCTGAGAGGCTGAGTGGGAGAGTGGCAGAG 411
 |||||
 QY 153 ValSerLysValProSerSerLeuSerGlnGlnInSerGlnInAspAlaIleTyrcIn 163
 |||||
 Db 412 ACTGCTGACATCTTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 427
 |||||
 RESULT 16
 AV012213 517 bp mRNA linear FRI 29 AUG 2001
 Locus Homo sapiens SDC SIGNB type IV isoform (G1209) mRNA, complete cds,
 alternatively spliced.
 DEFINITION
 AV012213.1 GI:15281696
 VERSION
 AV012213.1
 KEYWORDS
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS
 1 (bases 1 to 547)
 Jimenez,F., Abujar,S.S., and Abujar,S.K.
 Extensive repertoire of membrane-bound and soluble dendritic
 cell-specific ICAM-3-capturing nonintegrin 1 (DC-SIGN) and DC-SIGN2
 isoforms. Inter-individual variation in expression of DC-SIGN
 transcripts
 J. Biol. Chem. 276 (25), 3119-3222 (2001)
 MEDLINE
 21414847
 PUBMED
 11347487
 REFERENCE
 2 (bases 1 to 547)
 Mummidi,S. and Abujar,S.K.
 TITLE
 Direct Submission

[illegible]

US-09-831-458A 12 (1-25) x AY042225 (1-1105)

	AY042225	1143 bp mRNA	Human
LOCUS	AY042225		
DEFINITION	Human sapiens SDG-SIGNA 1 type I isoform (CD209) mRNA.		
ACCESSION	AY042225		
VERSION	AY042225.1	GI:15261080	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae; Homininae; Homo; Homo sapiens.		
AUTHORS	Mummid, S., Catalan, G., Lam, L., Hoeltje, A., Tejedor, J., Jimenez, F., Ahuja, S.S., and Ahuja, S.K.		
TITLE	Extensive repertoire of membrane-bound and soluble cell-specific TCM-1-grabbing integrin 1 (CD209) isoforms: Inter-individual variation in expression transcripts		
JOURNAL	J Biol Chem. 275 (45): 43396-4322 (2001)		
PUBLISHED	21413647		
REFERENCE	11337487		
AUTHORS	Mummid, S. and Ahuja, S.K.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-JUN-2001) Department of Medicine, S Veterans Health Care System and University of Texas Center, 7703, Floyd Curl Drive, San Antonio, TX 7		
FEATURES	Location/Qualifiers		
source	1..1143		
gene	/organism "Homo sapiens"		
CD209	/db_xref="taxon:9606"		
CD209	/chromosome "19"		
CD209	/map="19p14.3"		
CD209	1..1143		
CD209	/gene "CD209"		
CD209	1..1143		
CD209	/note="similar to the nucleotide sequence Genbank Accession Number AF208812; diff /codon_start 1		
CD209	/protein_id "AAK91850.1"		
CD209	/db_xref="gi:15261083"		
CD209	/translation "MSASAEPLQDLDLLEELKGLGFT QDSRDAIYONITDI KAAVGELSEKSIOEITQETKA LR LKAAGVLPERSKLQEIYGEILW KAAGVLPERSM EKSKOPIYGLTDLTKAAGVLPERSKOEDIVGLTHKAA NTLQAKAERLCPTPWTFPFGNGCTPSNSQSNMHISTE NLQDSSRSNRTWMISDI NQSGTWMDVGSFLISR EGVVKKSAEDV VLVKKEEDVAFV FSGNWGNDFNNFVFWCKSSASVSPNEDEPT SFVA		

BASE COUNT 309 a 296 c 440 g 208 t

ORIGIN

QY 98 LeuProGluLysSerLysLeuGlnGluIleTyrGlnIleuThrAryGluLysAlaAla 117
 DB 340 CTTCAGAGAAATCTAAGCTGACAGATCTACAGAGAGATGAGCTGAGAAATGTA 399

QY 118 ValGlyIleuLeuProGluLysSerLysLeuGlnGluIleTyrGlnIleuThr 135
 DB 400 GTGGGTGAGCTTCAGAGAAATCTAAGCTGACAGATCTACAGAGAGCTGACC 453

RESULT 24
 AX342003 1212 bp DNA linear PAT 10-JAN-2002
 LOCUS Sequence 9 from Patent W00196603.
 ACCESSION AX342003
 VERSION AX342003.1 GI:18137981
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo

REFERENCE
 1 Messier, W. and Sikela, J.M.
 Methods to identify polynucleotide and polypeptide sequences which
 may be associated with physiological and medical conditions
 Patent: W0 0196603-A 9 20-DEC-2001;
 Evolutionary Genomics, LLC (US)

FEATURES
 Source
 1. .1212
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 314 a 321 c 348 g 229 t
 ORIGIN

Alignment Scores:
 Pred. No.: 7.78e-72 Length: 1212
 Score: 78.00 Matches: 78
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 24.00% Indels: 0
 DB: 6 Gaps: 0

US-09-831-458a-12 (1-325) x AX342003 (1-1212)

QY 58 GlnAspAlaIleTyrGlnAsnIleuThrGlnIleuLysAlaAlaValGlyIleuSerGlu 77
 DB 220 CAAGACGCGATCTACAGAAATCTAAGCTGACAGATCTGAGCTGAGAAATGTA 219

QY 78 LysSerLysLeuGlnGluIleTyrGlnIleuThrGlnIleuLysAlaAlaValGlyIleu 97
 DB 280 AATCTCAAGCTTCAGAGAAATCTAAGCTGACAGATCTGAGCTGAGAAATGTA 349

QY 98 LeuProGluLysSerLysLeuGlnGluIleTyrGlnIleuThrAryGluLysAlaAla 117
 DB 340 CTTCAGAGAAATCTAAGCTGACAGATCTACAGAGAGATGAGCTGAGAAATGTA 399

QY 118 ValGlyIleuLeuProGluLysSerLysLeuGlnGluIleTyrGlnIleuThr 135
 DB 400 GTGGGTGAGCTTCAGAGAAATCTAAGCTGACAGAGATCTACAGAGAGCTGACC 453

RESULT 25
 AX039403 1215 bp DNA linear PAT 18-NOV-2000
 LOCUS Sequence 1 from Patent W00063251.
 ACCESSION AX039403
 VERSION AX039403.1 GI:11229473
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo

REFERENCE
 1 Fildner, C.G., Gelfond, R., van Kooijck, Y., and Torensmas, P.
 Composition and method for modulating dendritic cell-cell intera
 ction

JOURNAL Patent: W0 0063251-A 1 26-NOV-2000;
 FEATURES Katholieke Universiteit Nijmegen (NL)
 Source location/Qualifiers
 1. .1215
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Human; DC-SIGN"

BASE COUNT 315 a 321 c 349 g 230 t
 ORIGIN

Alignment Scores:
 Pred. No.: 7.79e-72 Length: 1215
 Score: 78.00 Matches: 78
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 24.00% Indels: 0
 DB: 6 Gaps: 0

US-09-831-458a-12 (1-325) x AX039403 (1-1215)

QY 58 GlnAspAlaIleTyrGlnAsnIleuThrGlnIleuLysAlaAlaValGlyIleuSerGlu 77
 DB 220 CAAGACGCGATCTACAGAAATCTAAGCTGACAGATCTGAGCTGAGAAATGTA 219

QY 78 LysSerLysLeuGlnGluIleTyrGlnIleuThrGlnIleuLysAlaAlaValGlyIleu 97
 DB 280 AATCTCAAGCTTCAGAGAAATCTAAGCTGACAGATCTGAGCTGAGAAATGTA 349

QY 98 LeuProGluLysSerLysLeuGlnGluIleTyrGlnIleuThrAryGluLysAlaAla 117
 DB 340 CTTCAGAGAAATCTAAGCTGACAGATCTACAGAGAGATGAGCTGAGAAATGTA 399

QY 118 ValGlyIleuLeuProGluLysSerLysLeuGlnGluIleTyrGlnIleuThr 135
 DB 400 GTGGGTGAGCTTCAGAGAAATCTAAGCTGACAGAGATCTACAGAGAGCTGACC 453

RESULT 26
 AY042221 1215 bp mRNA linear
 LOCUS Homo sapiens mdc-STONIA type 1 isoform (CD209) mRNA
 DEFINITION alternatively spliced.
 ACCESSION AY042221
 VERSION AY042221.1 GI:15291872
 KEYWORDS
 ORGANISM Homo sapiens.
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea;
 Mammidi; S. Catarrhini; Hominoidea; Hominidae;
 Hominini; S. Hominini; S. Hominini; S. Hominini;
 Jimenez, F., Ahuja, S.S., and Ahuja, S.K.
 Extensive repertoire of membrane-bound and soluble
 cell specific DCAM-3-grabbing neuropeptide 1 (CD-209)
 isoforms. Inter-individual variation in expression
 transcripts

JOURNAL J. Biol. Chem. 276 (35), 43395-43412 (2001)
 MEDLINE 2143847
 PUBMED 11337487

REFERENCE
 2 (bases 1 to 1215)
 Mummidi, S., and Ahuja, S.K.
 Direct Submission
 Submitted (18-JUN-2001) Department of Medicine, St
 Veterans Health Care System and University of Texas
 Center, 7703, Floyd Curl Drive, San Antonio, TX 782

FEATURES
 Source
 1. .1215
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19p13.3"
 1. .1215
 /gene="CD209"
 1. .1215

FEATURES	NEW YORK UNIVERSITY (US)	KATHOLIEKE UNIVERSITEIT NIJMEGEN (NL)
SOURCE	location/Qualifiers	
	1..1312	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
CDS	42..1256	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CNO12247.1"	
	/db_xref="GI:17049084"	
BASP ORIGIN	336 a 348 c 372 g 256 t	
Alignment Scores:		
Prod. No.:	8_24e-72	length: 1312
Score:	78.00	Matches: 78
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	24.00%	Indels: 0
DB:	6	Gaps: 0
US-09-841-458A-12 (1-325) x AX287103 (1-1312)		
QY	58 GlnAspAlaIleTyrGlnAsnLeuThrGlnLeuLysAlaAlaValGlyLeuSerGln	77
DB	261 CAAGACGCGATCTACGACGAACCTGACCTCAAGTTAAATATGATGAGGCTGACAG	320
QY	78 LysSerLysLeuIndIleTyrGlnGlnLeuThrGlnLeuLysAlaAlaValGlyLeu	97
DB	321 AATATCAAGCTGACGAGAGATCTACACGAGAGCTGACCGACGCTGAAGGCTGAGTGG	380
QY	98 LeuProGlnLysSerLysLeuGlnGlnIleTyrGlnGlnLeuThrArgLeuLysAlaAla	117
DB	381 CTTCGACACGAAATCTAAGCTCCACGAGATCTACACGAGCTGACCGCTGAAGGCTGGA	440
QY	118 ValGlyLeuLeuProGlnLysSerLysLeuIndIleTyrGlnGlnLeuThr	135
DB	441 GTGGGTAGCTTCAGAGAAATCTAAGCTTGACGAGATCTACGAGGCTGACG	494
RESULT 29		
LOCUS	HOMLECTINC	1312 bp mRNA linear PRE 15-JUN-2000
DEFINITION	Homo sapiens membrane-associated lectin in type-0 mRNA, complete cds.	
ACCESSION	M98457	
VERSION	M98457.1 GI:187115	
KEYWORDS		
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:	
	Mammalia: Eutheria: Primates: Catarrhini: Hominidae: Homo	
REFERENCE	1 (bases 1 to 1312)	
AUTHORS	Curtis,B.M., Scharnowski,S. and Watson,A.J.	
TITLE	Sequence and expression of a membrane-associated C-type lectin that	
	exhibits CD4-independent binding of human immunodeficiency virus	
	envelope glycoprotein gp120	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 89 (17), 8356-8360 (1992)	
MEDLINE	92390446	
PubMed	1518869	
FEATURES		
source	Location/Qualifiers	
	1..1312	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/tissue.type="placenta"	
CDS	42..1256	
	/codon_start=1	

	/product= "unibrain-associated lectin 133 /protein_id "AA077072.1" /db_xref="GI:8572943"
	/translat from "MSDSKPPHLLACVALLLEEDLRSLGH- LQLSFTLLAILLVSKSPSSISQSPQALYNLLIIL- ELTGKAAYGVEIPERSKIIEYGEITPIKAAYVELPEPSK- PEPSKOMETLYOETPIKAAYVELPEPSKOGLYEITPK- LTRIKAAYGEPERPSKAEIYGEITPIKAAYVELPEPSK- WHDSTLTACKEWADLVIVKSAIDFNEVLQDSSNFFIWM- ILPSPCKWNNKEENNVFEIEMDEFSGNSMDKCNLAFL- SPAPATPNPIYA"
BASE COUNT	336 a 348 c 372 g 256 t
ALIGNMENT SCORE:	
Prod No.:	8_24a-72 Length: 1312
Score:	78.00 Matches: 78
Percent Similarity:	100.00% Conserved: 0
Best local Similarity:	100.00% Mismatches: 0
Query Match:	24.00% Indels: 0
Df:	Gaps: 0
US-09-831-458A-12 (1-325) x H08LEPTING (1-1312)	
Oy	58 GlnAspAlalIeTyrGlnAsnArgThrGlnLeuysAlaAlaValcTyg-
Dd	261 CAATAAGCGAATCAACAGAGAATTAAGATGAGCTAACGGATGAGT
Oy	78 LysSerLysLeuGlnClnuIlleTyrGlnClnuIleThrGlnLeuysAlaAla
Dd	321 AAATCCAAATTCACAGCATCTACCAATCAATCACCAACTGAACCGCTG
Oy	98 LeuProGluLysSerLysLeuGlnClnuIleTyrGlnClnuIleThrArgArg
Dd	381 CTTCACAGCAAACTAAGCTGAGAGCATCTACAGAGAGCTGAGCCGCTG
Oy	118 ValcIglnLeuProGluLysSerLysLeuGlnClnuIleTyrGlnClnuIle
Dd	441 GTGGGTGAGCTTCAGACGAAATTAACTGTCAGAGATCTACGAGCT
RESULT 30	
LOCUS	AY042229 1115 bp mRNA Eukaryote
DEFINITION	Homo sapiens mbc StnIB type I isoform (CH299) gene
ACCESSION	AY042229
VERSION	AY042229.1 GI:15281088
KEYWORDS	
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; 1 (bases 1 to 1315) Mumudi,S., Gatano,G., Lam,D., Hoefle,A., Tolosa, Jimenez,F., Abuja,S.S. and Abuja,S.K. Extensive repertoire of membrane-bound and solu- cell-specific TCM-4-binding polypeptide 1 (BC2) isoforms, inter individual variation in expressed transcripts J Biol Chem 276 (45): 4219-4222 (2001)
JOURNAL	MEDLINE PubMed 1137487
REFERENCE	2 (bases 1 to 1315) Mumudi,S. and Abuja,S.K. Direct Submission Submitted (18-JUN-2001) Department of Medicine, Veterans Health Care System and University of Texas Center, 7703, Floyd Curl Drive, San Antonio, TX
TITLE	Location/Qualifiers 1..1315 /organism="Homo sapiens" /db_xref="taxon:9606" /chr_chromosome="19"
AUTHORS	
FEATURES	
Source	

FLV
LYC
GRI
YCH
ORN
SEN
CHI

```

gene
  /map="1p13.3"
  1..1315
  /gene="CD209"
misc_feature
  1..126
  /gene="CD209"
  /note="putative open reading frame; encodes truncated protein tdc-sicn18"
  /evidence="not_experimental"
  101..1315
CDS
  /gene="CD209"
  /note="Similar to the nucleotide sequence deposited in GenBank Accession Number AF290886, alternative splicing"
  /codon_start=1
  /product="MPC-SIGNB type I isoform"
  /protein_id="AAK91854.1"
  /db_xref="GI:15281090"
  /translation="MASAPVSTSTSTSHSEFQKSLSPGTPGYSKSLAGLGHPIV
  LLLSPFLLAGLVQVSKVPSISQBSRODAIYONLTOLKAAGELSEKSLQELIYQ
  ELTOLKAAGVLPKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQELTRIKAAVGL
  PEKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQ
  LTRIKAAVGLPEKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQ
  WHDSTTACKKVAQIVYIKSAEONFLQSSKSKRTWMLSLNBDTMOVDSIP
  LLPSFKQYWNKRGEPNNVGEFCARFSGNMMNCKNLAKFWIKKSAASCPDEOFL
  SPAPATPMPPPA"
BASE COUNT      332 a      345 c      346 j      252 t
ORIGIN
Alignment Scores:
  pred. No.:      8.25e-72      Length:      1315
  Score:          78.00      Matches:      78
  Percent Similarity: 100.00%      Conservative: 0
  Best local Similarity: 100.00%      Mismatches: 0
  Query Match:      24.00%      Indels:      0
  DE:              9      Gaps:      0
US-09-831-458a-12 (1-325) x AF290886 (1-1315)
CY 58 GlnAspAlaIleTyrGlnAsnLeuThrGlnLeuLysAlaIaValGlyLeuSerGln 77
14 320 CAAGACGCGATCTACCGAAGCCGAGCCAGGCTTAAAGCTGAGTGGAGCTCTCAGAG 379
CY 78 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 97
Db 380 AAATTCAGATCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAGAG 439
CY 98 LysProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
Db 440 GTTCCAGAGAACTAAGCTCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAGAG 499
CY 118 ValGlyGlyLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGln 135
Db 500 GTTCTGAGATCTTCCAGAGAAATCTAGAGATCTAGAGATCTAGAGATCTAGAG 553
RESULT 31
AF290886 4266 bp mRNA linear PRI 30 MAR 2001
DEFINITION Homo sapiens DC-SIGN mRNA, complete cds.
VERSION AF290886
KEYWORDS AF290886.1 GI:13324467
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4266)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Vilchev, S.J., Ellering, J.B.G., Martin, M.P., Wu, L., Martin, T.P.,
Vilchev, N., Knuckle, P.A., KewalRamani, V.N., van Kooijck, Y. and
Carroll, M.
A dendritic cell-specific intercellular adhesion molecule-
4-expressing nonintegrin (DC-SIGN)-related protein is highly
expressed on human liver sinusoidal endothelial cells and promotes
HIV-1 infection

```

```

JOURNAL J. Exp. Med. 193 (9), 671-676 (2001)
MEDLINE 21157496
PUBMED 11257134
REFERENCE 2 (bases 1 to 4266)
AUTHORS Hashitrova, A.A.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2000) ICD, NC1-FCR, P.O. Box B, Frederick, MD
21702, USA
FEATURES
  source
    Location/Qualifiers
      1..4266
      organism="Homo sapiens"
      db_xref="taxon:9606"
      chromosome="19"
      /map="1p13.3-p13.2"
      10..1224
      /codon_start=1
      /product="DC-SIGN"
      /protein_id="AAK20997.1"
      /db_xref="GI:13334468"
      /translation="MASAPVSTSTSHSEFQKSLSPGTPGYSKSLAGLGHPIV
      LLLSPFLLAGLVQVSKVPSISQBSRODAIYONLTOLKAAGELSEKSLQELIYQ
      ELTOLKAAGVLPKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQELTRIKAAVGL
      PEKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQ
      LTRIKAAVGLPEKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQELTRIKAAVGLPEKSKLOETIYQ
      WHDSTTACKKVAQIVYIKSAEONFLQSSKSKRTWMLSLNBDTMOVDSIP
      LLPSFKQYWNKRGEPNNVGEFCARFSGNMMNCKNLAKFWIKKSAASCPDEOFL
      SPAPATPMPPPA"
BASE COUNT      906 a      1170 c      995 g      1195 t
ORIGIN
Alignment Scores:
  pred. No.:      1.93e-71      Length:      4266
  Score:          78.00      Matches:      78
  Percent Similarity: 100.00%      Conservative: 0
  Best local Similarity: 100.00%      Mismatches: 0
  Query Match:      24.00%      Indels:      0
  DE:              9      Gaps:      0
US-09-831-458a-12 (1-325) x AF290886 (1-4266)
CY 58 GlnAspAlaIleTyrGlnAsnLeuThrGlnLeuLysAlaIaValGlyLeuSerGln 77
Db 739 CAAGACGCGATCTACCGAAGCCGAGCCAGGCTTAAAGCTGAGTGGAGCTCTCAGAG 288
CY 78 LysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 97
Db 289 AAATTCAGATCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAGAG 348
CY 98 LysProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 117
Db 349 GTTCCAGAGAACTAAGCTCTAGAGATCTAGAGATCTAGAGATCTAGAGATCTAGAG 408
CY 118 ValGlyGlyLeuProGlnLysSerLysLeuGlnGlnGlnGlnGlnGlnGlnGln 135
Db 409 GTTCTGAGATCTTCCAGAGAAATCTAGAGATCTAGAGATCTAGAGATCTAGAG 462
RESULT 32
AF290479 4424 bp DNA linear PRI 17 SEP 2000
DEFINITION Homo sapiens probable mannose-binding C-type lectin DC-SIGN gene.
VERSION AF290479
KEYWORDS AF290479.1 GI:10179609
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 4424)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Sollweh, E.J., Burton, K. and Townsend, J.
DC-SIGN: a related gene, DC-SIGN2, and CD23 form a cluster on 1p13
J. Immunol. 165 (6), 2937-2942 (2000)
20432267

```


FKQYWNKEPPNNVGEEDCAEFSGNGWMDKCNLAKFWICKKSAASCSRDDEOFLSPAP
ATNPPPA*
BASE COUNT 125 a 142 c 136 g 104 t
ORIGIN

Alignment Scores:

Prod. No.: 2.32e-12 Length: 507
Score: 21.00 Matches: 21
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.46% Indels: 0
DB: 9 Gaps: 0

US-09-831-458a-12 (1-325) x AY042224 (1-507)

UY 253 TRPMTGILYLeuSerASPLeuAsnGlnGluGlyThrIrpClnTrpValASPCLYScIPro 272
|||||
DB 245 TGGATGGGACTTTCATATCTAATCGAGTAATGACATGGCAATGGSTGGACGCTTACCT 294
UY 273 Leu 273
|||
DB 295 CTG 297

Search completed: December 7, 2002, 15:43:08
Job time : 3068 secs

•
•
•

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 7, 2002, 11:29:36 : Search time 262 seconds

(without alignments)
2793.509 Million cell updates/sec

Title: us-09-831-458a-12

Sequence: 1 MSKSKPRVWQLTLLGATGSH PRVWVWVWVKKPAPCPDE 325

Scoring table:

OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 1125939159 residues

Word size: 5

Total number of hits satisfying chosen parameters: 291269

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n model -DBY=XLB
-O/Cgna2_1/USP70_Spool/US0831458/runtat_05122002_103417_16781/3FP_query.fasta_1_519
-DB=N_Geneseq_101002 -OFMT=lastap -SUFFIX=oli.rng -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TPANS=bimandm.cdi
-LIST=45 -DOCCALIGN=200 -THR_SCORE=quality -THR_MINLEN=45 -MEME=LOCAL
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=0509631458_wgn_1_1_0_runtat_05122002_103417_16781 -RCPU 6 -RCPU=3
-NO_XLPHY -NO_MMAP -LARGEMEMORY -NFG_STOPPS=0 -MAIT -LONQ_OG -PRV_TTMPOUT=120
-MARR_TTMPOUT=10 -THREPOUS=1 -XGAPOP=60 -XGAPEXT=60 -FSGAPOP=6 -FSGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

N_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Desc
1	325	100.0	1542	21	AAA27055	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
2	168	51.7	1643	20	AAV84361	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
3	168	51.7	1643	22	AAV5469	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
4	168	51.7	1643	24	AHL41993	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
5	97	29.8	792	22	AAH8239	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
6	97	29.8	2991	22	AAH16026	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
7	84	25.8	553	22	AAH67709	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
8	78	24.0	1212	21	AAV41226	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
9	78	24.0	1215	21	AAH8740	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
10	78	24.0	1215	21	AAH65383	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
11	78	24.0	1312	14	AAU56648	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
12	78	24.0	1312	22	AAV5461	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
13	55	16.9	433	22	AAH99465	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
14	51	15.7	592	22	AAH60904	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
15	51	15.7	592	22	AAH28884	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
16	51	15.7	592	22	AAH69194	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
17	51	15.7	592	22	AAH35083	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
18	51	15.7	592	22	AAH40799	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
19	51	15.7	592	24	AAH9548	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
20	50	15.4	152	22	AAH43421	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
21	50	15.4	152	22	AAH48762	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
22	50	15.4	152	22	AAH21864	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
23	50	15.4	152	22	AAH48027	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
24	50	15.4	152	22	AAH38856	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
25	50	15.4	152	24	AAH21963	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
26	43	13.2	1212	24	AAH41224	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
27	35	10.8	1212	24	AAV81225	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
28	29	9	265	20	AAH89716	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
29	29	9	606	24	AAH20341	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
30	30	9	1183	18	AAH60314	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
31	31	9	1516	22	AAH42221	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
32	32	9	2853	18	AAH22559	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
33	34	9	2648	18	AAH12558	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
34	34	9	2755	24	AAH12899	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
35	35	9	2757	22	AAH41177	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
36	36	9	2757	24	AAH49160	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
37	37	9	2757	24	AAH40585	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
38	38	9	2781	24	AAU49159	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
39	39	9	2888	24	AAU49162	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
40	40	9	2892	24	AAU49161	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
41	41	9	3512	24	AAU49163	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
42	42	9	4418	22	AAH2159	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
43	43	9	16424	22	AAH68448	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
44	44	9	268	22	AAH2567	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
45	45	9	14005	24	AAH98600	Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss

ALIGNMENTS

RESULT 1

AAA27055 standard; cDNA: 1542 bp.

AAA27055: 22-AUG-2000 (first entry)

Human cell surface receptor protein cDNA sequence #12.

Human, HCSR, cytosolic; antiarrhythmic; antiinflammatory; immunosuppressive; antiarteriosclerotic; antibacterial; neuroprotective; nootropic; anticonvulsant; cancer; leuk melanoma; rheumatoid arthritis; asthma; atherosclerosis; Alzheimer's diseases; multiple sclerosis; epilepsy; ss
Homo sapiens.
XX
OS
XX

Key	Location/qualifiers
FT CDS	269..1246
FT FT	/*tag= a
FT FT	/product= "HCSR-12"
LN	W0200028032-A2.
FD	18-MAY-2000.
FE	12-NOV-1999; 99MO-US26742.
FE	12-NOV-1998; 98US-0191280.
PR	07-SEP-1998; 98US-0206447.
PR	08-MAR-1999; 99US-0123404.
PA	(INCY-) INCYTE PHARM INC.
XX	
P1	Tanq YT, Corley NC, Guebler KJ, Yue H, Baughn MB, Lal P;
P1	Hillman JL, Bandman O, Azimzal Y, Au-Young J;
XX	
DE	WPI: 2000-376546/32.
DR	P-PSDB; AAN94345.
XX	
P1	New human cell surface receptor protein and polynucleotide useful for
P1	diagnosis, prevention and treatment of cancer, immune disorders,
P1	infection and neuronal disorders -
XX	
PS	Claim 9; page 36-97; 97pp; English.
XX	
CC	The present sequence encodes a novel human cell surface receptor protein
CC	(HCSR) designated HCSR-12. The nucleotide sequence was identified in
CC	(HCSR) Clone 3144966 from the cDNA library SPUNOT09, which was made
CC	from RNA isolated from diseased spleen tissue. A number of cDNA clones
CC	were used to assemble the consensus sequence. BLAST analysis showed that
CC	the sequence is homologous to non-CD4 glycoprotein gp120 receptor
CC	AAR32188. HCSR and its antagonist are useful for preventing or treating
CC	disorders associated with decreased or increased expression of activity
CC	of HCSR. Such disorders include cancers such as leukemia and melanoma,
CC	immune disorders such as rheumatoid arthritis, asthma and
CC	atherosclerosis, bacterial and parasitic infections and neuronal
CC	disorders such as Alzheimer's disease, multiple sclerosis and
CC	epilepsy. Polynucleotides encoding HCRPs may be used as hybridisation
CC	probes to diagnose these conditions. Anti-HCSR antibodies may be used
CC	as antagonists, as a targeting or delivery mechanism for binding
CC	pharmaceutical agents into contact with cells or tissues expressing
CC	HCSR and for diagnosis of HCSR-related disorders. HCSR and its
CC	catalytic or immunogenic fragments are useful for drug screening using
CC	libraries of compounds.

SV	Sequence 1542 BF: 385 A; 417 G; 408 G; 335 T; 0 other:						
<hr/>							
Alignment Scores:							
Pred. No.:	1,39e-313	Length:	1542				
Scores:	325..00	Matches:	325				
Percent Similarity:	100.00%	Conservative:	0				
Best Local Similarity:	100.00%	Mismatches:	0				
Query Match:	100.00%	Indels:	0				
DB:	21	Gaps:	0				
<hr/>							
05-09-831-458A-12 (1-325) x AA027055 (1-1542)							
CY	1 MetSerAspSerLeuGluProArgValClnIcInuGlyLeuDeuCIySLacGlyHis	20					
DB	259 ATGAGTACTCAATGAACAACTGGTCACCATGTGGTGCGCTGTAAGTCTTTGGCAT	325					
CY	21 CTGATATGVALLeuSerInuLeuLysSerThomelLeuAla3GlyValLeuValAlle	40					
DB	329 GAGGAGCTGGTGTGTAAATCTCTTCTTTATAATATTGGTATGGTATGTG	385					
CY	41 LeuValClnValSerLysValProSerSerLeuSerGlnIcInuSerGlnIcAspaLa	60					
DB	389 GTTTTTCAAATTTTCAAAGTTTTTAAATTCTTAAATTAAGTAACAATATAATAG	444					

[illegible]

10	AAV84361	Scd44orf; cDNA; 1043 bp.
XX	AC	AAV84361;
XX	XX	30-MAR-1999 (first entry)
DT	XX	Human liver cDNA clone Hp01347.
XX	DE	Transmembrane protein; HPO1347; human; lectin; receptor; liver; ds
XX	KW	Homo sapiens.
XX	OS	Key
XX	XX	25,915
FT	CDs	/ftad = d
FT	FT	location/qualifiers

US-09-831-458A-12 (1-325) x AAH07709 (1-553) English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification, where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides, and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialized methods. AAH03166 to AAH15628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to

XX AAH95893 represent human amino acid sequences; and AAH15629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

XX Sequence 2991 BP: 778 A; 773 G; 827 G; 613 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 1,986-86 Length: 2991

XX Score: 97.00 Matches: 124

XX Percent Similarity: 98.41% Conservative: 0

XX Best Local Similarity: 98.41% Mismatches: 1

XX Query Match: 29,85% Indels: 2

XX DB: 22 Gaps: 0

US-09-831-458A-12 (1-325) x AAH07709 (1-553)

XX 44 ValSerLysValProSerSerLeuSerGlnGlnGlnSerGlnAspAlaIleTyrIle 63

XX 53 GTCGTCACACCTCCAGCTCCCTAACTAGGAACTAATCGACACACACGACATCTACG 112

XX 64 AsnLeuThrGlnLeuLysAlaLysValGlnGlnGlnSerGlnLysSerLysGlnGln 83

XX 113 AACCTACCTACCTTAAAGTTAC-ATGAGGCTAGAGCTCTGACAGGAAATCCAACTGCAAGA 171

XX 84 ValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 102

XX 172 GATCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTAC 241

XX 103 GlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 123

XX 232 GCTGACGACATCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTAC 291

XX 123 LysSerLysLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 143

XX 292 GAAATCTAACTTAAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 351

XX 143 ValLeuProGlnLysSerLysLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 163

XX 452 GTTACACAGCAAAATCCAAAGCTGACAGCAATCTACAGGAGCTGACAGGAGCTGACAGG 411

XX 163 ValGlnGlnLeuPro 168

XX 412 AGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 427

XX RESULT 7

XX AAH07709

XX AAH07709 standard: cDNA: 553 BP

XX 26-1001 (first entry)

XX

DE Human cDNA clone (5' primer) SEQ ID NO:4544.

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

XX Homo sapiens

US-09-831-458A-12 (1-325) x AAH07709 (1-553)

XX Sequence 553 BP: 149 A; 144 G; 151 G; 104 T; 5 other;

XX Alignment Scores:

XX Pred. No.: 3,486-74 Length: 553

XX Score: 84.00 Matches: 84

XX Percent Similarity: 100.00% Conservative: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 29,85% Indels: 0

XX DB: 22 Gaps: 0

XX 16 GTCGTCACACCTCCAGCTCCCTAACTAGGAACTAATCGACACACACGACATCTACG 35

XX 183 GTCGTCACACCTCCAGCTCCCTAACTAGGAACTAATCGACACACACGACATCTACG 242

XX 26 ValLeuValAlaIleLeuValGlnValSerLysSerLeuSerGlnGlnGlnGlnGlnGln 55

XX 113 AACCTACCTACCTTAAAGTTAC-ATGAGGCTAGAGCTCTGACAGGAAATCCAACTGCAAGA 171

XX Sequence 592 BP, 154 A, 149 C, 168 G, 121 T, 0 other:
 Alignment Scores:
 Pred. No.: 2,56e-41 Length: 592
 Score: 51.00 Matches: 51
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15,69% Indels: 0
 DB: 22 Gaps: 0
 US-09-831-458a-12 (1-525) x AAK35083 (1-592)
 Q7 188 G1AARLEAGYSAARHISCTYSPTCTYSASTTETLPHLEPLHGLCYASCTSTPTTHe 207
 DB 149 GAACGCTGTGGCCGACCTGTCTTGAAGGACATTTCTTGAAGGAAACCTTACTTG 208
 Q7 208 MetSerAsnSerGlnArgAsnGlnPhtAspSerValThrAlaGlyGlnGlnValAla 227
 DB 209 ATGTCTAACTCCGACGCACTGCTGCTAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 274
 Q7 228 G1AARLEAGYSAARHISCTYSPTCTYSASTTETLPHLEPLHGLCYASCTSTPTTHe 238
 DB 269 CAAGCTGTCTGTAATCAAAACTGCTGAGGACAG 301
 RESULT 17
 AAK35083
 AAK35083 standard: DNA: 592 BP.
 AAK35083:
 05-NOV-2001 (first entry)
 XX
 XX Human bone marrow expressed single exon probe: SEQ ID NO: 9646.
 XX
 XX Human bone marrow expressed exon: gene expression analysis: probe:
 XX microarray: cancer; leukaemia; lymphoma; myeloma; ss.
 XX
 XX Homo sapiens
 XX
 XX W0200157276-A2
 XX
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001: 2001WO-US00668.
 XX
 XX 04-FEB-2000: 2000US-0180312.
 XX 26-MAY-2000: 2000US-0207456.
 XX 30-JUN-2000: 2000US-060408.
 XX 03-AUG-2000: 2000US-0632466.
 XX 21-SEP-2000: 2000US-0234687.
 XX 27-SEP-2000: 2000US-0236359.
 XX 01-DEC-2000: 2000US-0621264.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DP:
 XX WPI: 2001-488900/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human bone marrow.
 XX
 XX Example 4: SEQ ID NO: 9646; 654bp + Sequence Listing: English.
 XX
 XX The present invention provides a number of single exon nucleic acid
 XX probes which are derived from genomic sequences expressed in the human
 XX bone marrow. They can be used to measure gene expression in bone marrow
 XX samples, which may enable the improved diagnosis and treatment of cancers
 XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
 XX the probes of the invention.
 XX
 XX Sequence 592 BP, 154 A, 149 C, 168 G, 121 T, 0 other:

Alignment Scores:
 Pred. No.: 2,56e-41 Length: 592
 Score: 51.00 Matches: 51
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15,69% Indels: 0
 DB: 22 Gaps: 0
 US-09-831-458a-12 (1-525) x AAK35083 (1-592)
 Q7 188 G1AARLEAGYSAARHISCTYSPTCTYSASTTETLPHLEPLHGLCYASCTSTPTTHe 207
 DB 149 GAACGCTGTGGCCGACCTGTCTTGAAGGACATTTCTTGAAGGAAACCTTACTTG 208
 Q7 208 MetSerAsnSerGlnArgAsnGlnPhtAspSerValThrAlaGlyGlnGlnValAla 227
 DB 209 ATGTCTAACTCCGACGCACTGCTGCTAATCTGCTGCTGCTGCTGCTGCTGCTGCTG 274
 Q7 228 G1AARLEAGYSAARHISCTYSPTCTYSASTTETLPHLEPLHGLCYASCTSTPTTHe 238
 DB 269 CAAGCTGTCTGTAATCAAAACTGCTGAGGACAG 301
 RESULT 18
 AAK40799
 AAK40799 standard: DNA: 592 BP.
 AAK40799:
 17-OCT-2001 (first entry)
 XX
 XX Probe #9485 used to measure gene expression in human placenta sample.
 XX
 XX Probe: microarray: human; placenta; antenatal diagnosis:
 XX genetic disorder; ss.
 XX
 XX Homo sapiens
 XX
 XX W0200157272-A2
 XX
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001: 2001WO-US00664.
 XX
 XX 04-FEB-2000: 2000US-0180312.
 XX 26-MAY-2000: 2000US-0207456.
 XX 30-JUN-2000: 2000US-060408.
 XX 03-AUG-2000: 2000US-0632466.
 XX 21-SEP-2000: 2000US-0234687.
 XX 27-SEP-2000: 2000US-0236359.
 XX 01-DEC-2000: 2000US-0621264.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DP:
 XX WPI: 2001-488900/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human placenta.
 XX
 XX Claim 25: SEQ ID NO 9485; 654bp; English.
 XX
 XX The present invention relates to single exon nucleic acid probes (SENP).
 XX The present sequence is one such probe. The probes are useful for
 XX producing a microarray for predicting, measuring and displaying gene
 XX expression in samples derived from human placenta. The probes are useful
 XX for antenatal diagnosis of human genetic disorders.
 XX
 XX Sequence 592 BP, 154 A, 149 C, 168 G, 121 T, 0 other:
 Alignment Scores:
 Pred. No.: 2,56e-41 Length: 592

FN W0200157277 A2.
 XX
 XX 09-AUG-2001
 XX
 XX 40-JAN-2001; 2001WO-0500669.
 XX
 XX 04-FEB-2000; 2000US-0180412
 XX 25-MAY-2000; 2000US-0207454
 XX 30-MAY-2000; 2000US-0608408
 XX 03-AUG-2000; 2000US-0642564
 XX 21-SEP-2000; 2000US-0234487
 XX 27-SEP-2000; 2000US-0234487
 XX 04-OCT-2000; 2000US-0024463.
 XX
 XX (MOLE-) MOLEBYLAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human fetal liver -
 XX
 XX Claim 4; SEQ ID NO 21726; 649pp; sequence listing; English.
 XX
 XX The invention relates to a single exon nucleic acid probe for
 XX measuring human gene expression in a sample derived from human fetal
 XX liver. The single exon nucleic acid probes may be used for predicting,
 XX measuring and displaying gene expression in samples derived from human
 XX fetal liver. The present sequence is a single exon nucleic acid
 XX probe of the invention.
 XX Note: The sequence data for this patent did not form part of the
 XX printed specification, but was obtained in electronic format directly
 XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 152 BP, 37 A, 48 C, 38 G, 23 T, 0 other.
 XX
 XX Alignment Scores:
 XX Pred. No: 7e-41 Length: 152
 XX Score: 50.00 Matches: 50
 XX Percent Similarity: 100.00% Conservativ: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 15,388 Indels: 0
 XX DB: 22 Gaps: 0
 XX
 XX US-09-831-458A-12 (1-325) x ABA73421 (1-152)
 XX
 XX 189 Arleuowfysarfhscysprrpoyasprftrphphocfncjyvasocystyrthwot 208
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 XX 4 CGGCTGTGGGCGGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGACG 62
 XX
 XX 209 Setasnsrslarqasnrtphtsaspservalthralacysglndluvalarqalacn 228
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 XX 63 TCTAAGTCCGACGCGAAGCTGTCACGTCGTCGACGCGCTGCGACGAGTAAGTGGCGCGAC 122
 XX
 XX 229 LeuvaiVala|||cys|thralac|uc|uc|n 238
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 XX 123 CTGCTGTAAATCAAACTGCTGAGGACAG 152
 XX
 XX RESULT 21
 XX ABA73421
 XX ID ABA73421 standard; DNA; 152 BP.
 XX
 XX ABA73421:
 XX
 XX 24-JAN-2002 (first entry)
 XX
 XX Probe #17218 for gene expression analysis in human heart cell sample.
 XX
 XX human; gene expression; heart; microarray; vascular system; probe;
 XX cardiovascular disease; hypertension; cardiac arrhythmia;
 XX congenital heart disease; ss.
 XX
 XX

CS Homo sapiens.
 XX
 XX W0200157277 A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001WO-0500669.
 XX
 XX 04-FEB-2000; 2000US-0180412
 XX 25-MAY-2000; 2000US-0207454
 XX 30-MAY-2000; 2000US-0608408
 XX 03-AUG-2000; 2000US-0642564
 XX 21-SEP-2000; 2000US-0234487
 XX 27-SEP-2000; 2000US-0234487
 XX 04-OCT-2000; 2000US-0024463.
 XX
 XX (MOLE-) MOLEBYLAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483447/52.
 XX
 XX Single exon nucleic acid probes for analyzing gene expression in human
 XX hearts -
 XX
 XX Claim 4; SEQ ID NO 17218; 549pp; English.
 XX
 XX The present invention relates to single exon nucleic acid probes for
 XX measuring human gene expression in a sample derived from human heart. The
 XX present sequence is one such probe. The probes may be used for
 XX predicting, measuring and displaying gene expression in samples derived
 XX from the human heart via microarrays. By measuring gene expression, the
 XX probes are useful for predicting, diagnosing, grading, staging,
 XX monitoring and proposing diseases of the human heart and vascular system
 XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 XX congenital heart disease.
 XX Note: The sequence data for this patent did not form part of the printed
 XX specification, but was obtained in electronic format directly from WIPO
 XX at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 152 BP, 37 A, 48 C, 38 G, 23 T, 0 other.
 XX
 XX Alignment Scores:
 XX Pred. No: 7e-41 Length: 152
 XX Score: 50.00 Matches: 50
 XX Percent Similarity: 100.00% Conservativ: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 15,388 Indels: 0
 XX DB: 22 Gaps: 0
 XX
 XX US-09-831-458A-12 (1-325) x ABA73421 (1-152)
 XX
 XX 189 Arleuowfysarfhscysprrpoyasprftrphphocfncjyvasocystyrthwot 208
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 XX 4 CGGCTGTGGGCGGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGACG 62
 XX
 XX 209 Setasnsrslarqasnrtphtsaspservalthralacysglndluvalarqalacn 228
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 XX 63 TCTAAGTCCGACGCGAAGCTGTCACGTCGTCGACGCGCTGCGACGAGTAAGTGGCGCGAC 122
 XX
 XX 229 LeuvaiVala|||cys|thralac|uc|uc|n 238
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 XX 123 CTGCTGTAAATCAAACTGCTGAGGACAG 152
 XX
 XX RESULT 22
 XX AAK21864
 XX ID AAK21864 standard; DNA; 152 BP.
 XX
 XX AAK21864:
 XX
 XX 05-NOV-2001 (first entry)
 XX
 XX Human brain expressed single exon probe SEQ ID NO: 21855.
 XX
 XX

```

XX  Human: brain expressed exon; gene expression analysis; probe;
KW  microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KM  epilepsy; cancer; ss.
XX
OS  Homo sapiens.
XX
PN  W0200157275-A2.
XX
PD  09-AUG-2001.
XX
PE  10-JAN-2001; 2001W0-US009667.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  10-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0612366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0246359.
PR  04-OCT-2000; 2000US-0024263.
XX
XX  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PA  Penn SG, Hanzel DK, Chen W, Rank DR;
PI  WPI; 2001-483446/52.
XX
DR  WPI; 2001-483446/52.
XX
PT  Single exon nucleic acid probes for analyzing gene expression in human
PS  brains.
XX
XX  Example 4; SEQ ID NO: 21855; 650bp. Sequence listing: English.
XX
CC  The present invention provides a number of single exon nucleic acid
CC  probes which are derived from genomic sequences expressed in the human
CC  brain. They can be used to measure gene expression in brain cell samples,
CC  which may enable the diagnosis and improved treatment of nervous system
CC  diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC  epilepsy and cancers. The present sequence is one of the probes of the
CC  invention.
XX
SO  Sequence 152 BP; 37 A; 48 C; 38 G; 29 T; 0 other;
XX
XX  Alignment Scores:
XX  Pred. No.: 7e-41 Length: 152
XX  Score: 50.00 Matches: 50
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 15,388 Indels: 0
XX  DB: 22 Gaps: 0
XX
US-09-831-458a-12 (1-325) x AAK21864 (1-152)
QY 189 ArgLeuGysArhHisCysProLysAspTyrPhePheGlnGlyAsnGlySylrPheMet 208
DB 3 GGRCTGGGCGGACGCTGCTCCAGAGCTGACATATTCCTCAAGCAAACTTACTCTCATG 62
QY 209 SerAsnSerGlnArgAsnTyrPheAspSerValIleAlaGlySjlnIleValAlaGln 228
DB 63 TCTAACTCCGAGCGGAACTGATATGATCTGGTATGATCTGGTATGATCTGGTATG 122
QY 229 LeuValValIleLeuThrAlaGlnGln 238
DB 123 CTGCTGTAATCAAACTGCTGAGGAGCAG 152
XX
RESULT 24
AAK48027
ID AAK48027 standard; DNA; 152 BP.
XX
AC AAK48027;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe; SEQ ID NO: 22584.
XX

```

```

XX  Human: bone marrow expressed exon; gene expression analysis;
KW  microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS  Homo sapiens.
XX
PN  W0200157276-A2.
XX
PD  09-AUG-2001.
XX
PE  30-JAN-2001; 2001W0-US009669.
XX
PR  04-FEB-2000; 2000US-0180312.
PR  26-MAY-2000; 2000US-0207456.
PR  10-JUN-2000; 2000US-0608408.
PR  03-AUG-2000; 2000US-0612366.
PR  21-SEP-2000; 2000US-0234687.
PR  27-SEP-2000; 2000US-0246359.
PR  04-OCT-2000; 2000US-0024263.
XX
XX  (MOLE-) MOLECULAR DYNAMICS INC.
XX
PA  Penn SG, Hanzel DK, Chen W, Rank DR;
PI  WPI; 2001-488900/53.
XX
DR  WPI; 2001-488900/53.
XX
PT  Human genome-derived single exon nucleic acid probes used
PS  analyzing gene expression in human bone marrow.
XX
XX  Example 4; SEQ ID NO: 22584; 650bp. Sequence listing: English.
XX
CC  The present invention provides a number of single exon nucleic acid
CC  probes which are derived from genomic sequences expressed in human
CC  bone marrow. They can be used to measure gene expression in brain
CC  samples, which may enable the improved diagnosis and treatment of
CC  diseases such as lymphoma, leukemia and myeloma. The present sequence
CC  is one of the probes of the invention.
XX
SO  Sequence 152 BP; 37 A; 48 C; 38 G; 29 T; 0 other;
XX
XX  Alignment Scores:
XX  Pred. No.: 7e-41 Length: 152
XX  Score: 50.00 Matches: 50
XX  Percent Similarity: 100.00% Conservative: 0
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 15,388 Indels: 0
XX  DB: 22 Gaps: 0
XX
US-09-831-458a-12 (1-325) x AAK48027 (1-152)
QY 189 ArgLeuGysArhHisCysProLysAspTyrPhePheGlnGlyAsnGlySylrPheMet 208
DB 3 GGRCTGGGCGGACGCTGCTCCAGAGCTGACATATTCCTCAAGCAAACTTACTCTCATG 62
QY 209 SerAsnSerGlnArgAsnTyrPheAspSerValIleAlaGlySjlnIleValAlaGln 228
DB 63 TCTAACTCCGAGCGGAACTGATATGATCTGGTATGATCTGGTATGATCTGGTATG 122
QY 229 LeuValValIleLeuThrAlaGlnGln 238
DB 123 CTGCTGTAATCAAACTGCTGAGGAGCAG 152
XX
RESULT 24
AAI53856
ID AAI53856 standard; DNA; 152 BP.
XX
AC AAI53856;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #22542 used to measure gene expression in human bone marrow;
XX  Probe, microarray, human, placenta, antenatal diagnosis;
XX

```

KM gene to disorder: SS
 XX
 US Homo sapiens
 XX
 PN WO200157272-A2
 XX
 XX
 PD 09-AUG-2001
 XX
 PF 30-JAN-2001: 2001WO-0500663
 XX
 PR 04-FEB-2000: 2000US-0180312
 PR 26-MAY-2000: 2000US-0207456
 PR 30-JUN-2000: 2000US-0608408
 PR 03-AUG-2000: 2000US-0632366
 PR 21-SEP-2000: 2000US-0234687
 PR 27-SEP-2000: 2000US-0234686
 PR 04-OCT-2000: 2000US-0024263
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn St, Hanzel PK, Chen W, Park DR
 XX
 DR WPI: 2001-488897/53
 XX
 PT Human genome derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 25: SEQ ID NO 22542; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC predicting a multiplicity for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 CC
 XX
 SQ Sequence 152 bp, 37 A; 48 C, 38 G, 23 T, 0 other;
 XX
 Alignment Scores:
 Pref. No.: 7e-41 Length: 152
 Score: 50.00 Matches: 50
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatch: 0
 Query Match: 15,388 Indels: 0
 Dbs: 22 Gaps: 0
 US-09-831-458A-12 (1-325) x AA153856 (1-152)
 YY 189 ArgLacGcATGhGcScFrcGgSAsTfPcAlPhcPhgGlnGlyAsnGcSTyDhaMet 208
 Db 3 CGCGTGTGCGGCTACTGGTAAAGGAGTGAATATGTCGAAAGCAATGTTACTTCATG 62
 YY 209 SetAsnSerGlnARgASnTrpHisAspSerValIluAlaGcGlnGlnValATGAlaGln 238
 Db 63 TCTAATTCGACGGGTAACGTCACGACGTCGTCGACGAGGCTGTCGAGGAGGCTGTCAG 122
 YY 229 GcGValValIleGlyThrAlaGlnGln 238
 Db 123 GTGTCGTATATCAAAAGCTGCTAGGAGGAG 152
 RESULT 25
 ID ABS21963 standard, DNA, 152 bp.
 XX
 AC ABS21963;
 XX
 DI 19-AUG-2002 (first entry)
 XX
 DE Human genome-derived single exon probe from lung SEQ ID NO 21954.
 KW Human ds; single exon probe; asthma, lung cancer, COPD, ILD;
 KW chronic obstructive pulmonary disease; interstitial lung diseases;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;

KM Human ds; single exon probe; asthma, lung cancer, COPD, ILD;
 KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karagane syndrome;
 KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KM primary ciliary dyskinesia; pulmonary hypertension;
 XX hyaline membrane disease; open reading frame, ORF.
 XX
 OS Homo sapiens.
 OS
 PN WO200160004-A2
 XX
 PD 15 NOV 2001
 XX
 PF 30 JAN 2001: 2001WO-0500663
 XX
 PR 04-FEB-2000: 2000US-180312P
 PR 26-MAY-2000: 2000US-0207456P
 PR 30-JUN-2000: 2000US-0608408
 PR 03-AUG-2000: 2000US-0632366
 PR 21-SEP-2000: 2000US-0234687P
 PR 27-SEP-2000: 2000US-0234686P
 PR 04-OCT-2000: 2000US-0024263
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn St, Hanzel PK, Chen W, Park DR;
 XX
 DR WPI: 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 4: Seq ID No 21954; 654pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA; and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC ESTs; a set of reading frames (ORFs). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberculous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Heremansky Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Kartagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: the sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

```
XX      Sequence 152 BP: 37 A; 48 G; 38 C; 29 T; 0 other.
SQ
Alignment Scores:
Pred. No.:      7e-41      Length:      152
Score:          50.00      Matches:      50
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      15.38%      Indels:      0
DE:              24      Gaps:      0
US-09-831-458a-12 (1-325) x ABA91224 (1-1212)
OY      189 ArgLeuGysArqHsYsPpPolysAsPTrPhePheGlnGlnYsncYsTyrrpPheMet 208
Db      3 CGCCCTGTCGCCGCACTGTCGCCAAGCACTGCAATCTTCTCCAGGAAAGCTGTACTTCATG 62
OY      209 SerAsnSerGlnArqAsnTrfHISAsrSeiValThrAlaGysGlnGlnValArqGAla3in 228
Db      63 TCTAAGTCCCAAGGCGAACTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
OY      229 LeuValValIleIysThrAlaGlnGln 238
Db      123 CTCTCTGTAATCAAAACTGCTGACGACGAC 152
RESULT 26
ABA91224
ID      ABA91224 standard; cDNA; 1212 BP.
XX
AC      ABA91224:
XX
DT      04-APR-2002 (first entry)
DE      chimpanzee DC-SIGN cDNA coding region.
XX
KW      DC-SIGN; chimpanzee; gene; genomics; evolution; AIDS; HIV-1; s3.
XX
OS      Pan troglodytes.
XX
PN      W0200196603-A2.
XX
PD      20-DEC-2001.
XX
PE      06-JUN-2001: 2001MO-US18310.
XX
PR      09-JUN-2000: 2000US-0591435.
XX
PA      (EVOL-) EVOLUTIONARY GENOMICS LLC.
XX
PI      Messier W, Sikela JM.
XX
PT      WPI: 2002-130744/17.
XX
DE      Determining human or non-human primate polynucleotide or polypeptide
PT      sequences associated with a physiological trait and have undergone
PT      evolutionary changes, for therapeutic use, involves using statistical
PT      methods -
XX
XX      Example 19, Fig 12, 14pp, English.
XX
CC      The present sequence is that of the coding region of the chimpanzee
CC      DC-SIGN gene. DC-SIGN is expressed on dendritic cells and is known
CC      to provide a mechanism for transport of HIV-1 virus to the lymph
CC      nodes. HIV-1 binds to the extracellular portion of DC-SIGN and
CC      infects the undifferentiated T cells in the lymph nodes via their
CC      CD4 proteins. This ultimately leads to compromise of the immune
CC      system and to full-blown AIDS. The invention comprises a
CC      comparative genomics approach to identify specific gene changes
CC      responsible for differences in functions and diseases distinguishing
CC      humans from non-humans, particularly primates including gorilla,
CC      orangutan and especially chimpanzee. The evolutionary significance
CC      of a nucleotide change is determined by the ratio of the
CC      non-synonymous substitution rate (KA) to the synonymous rate (KS)
```

```
CC      of the nucleotide sequence. Polynucleotide and polypeptide
CC      sequences corresponding to evolved traits may be relevant to
CC      diseases or conditions such as unique or enhanced human
CC      functions, longer human life spans, susceptibility or resistance
CC      to disease, including AIDS and cancer, and aesthetic traits
CC      such as hair growth. KA/KS ratios for chimpanzee/human, human/
CC      chimpanzee/gorilla DC-SIGN cDNA sequences are 1.3, 0.87 and
CC      respectively. It is theorized that chimpanzee resistance
CC      progression to full-blown AIDS may be due in part to the
CC      HIV-1 to bind to chimpanzee DC-SIGN for transport to the
CC      nodes. After determining the 3-dimensional structure of
CC      a rational drug design approach can be used to mimic the
CC      of chimpanzee DC-SIGN without interfering with the normal
CC      of human DC-SIGN.
XX
SQ      Sequence 1212 BP: 315 A; 216 G; 459 C; 230 T; 1 other.
OY      61 lETyrgInAsnLeuThrGlnLeuGysAlaValaGlyGlnLeuSerGln 1212
Db      229 ALCIAGTGAAGACGACCCACTTAAGACGACGACGACGACGACGACGACGACGACGACGACGAC 1212
OY      81 LeuGlnGlnIleTyrgInGlnGlnLeuThrGlnLeuGysAlaValaGlyGln 1212
Db      289 CTCTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1212
OY      101 LysSerLys 103
Db      349 AATCTTAAG 357
RESULT 27
ABA91225
ID      ABA91225 standard; cDNA; 1212 BP.
XX
AC      ABA91225:
XX
DT      04-APR-2002 (first entry)
DE      Gorilla DC-SIGN cDNA coding region.
XX
KW      DC-SIGN; gorilla; gene; genomics; evolution; AIDS; HIV-1.
XX
OS      Gorilla gorilla.
XX
PN      W0200196603-A2.
XX
PD      20-DEC-2001.
XX
PE      06-JUN-2001: 2001MO-US18310.
XX
PR      09-JUN-2000: 2000US-0591435.
XX
PA      (EVOL-) EVOLUTIONARY GENOMICS LLC.
XX
PI      Messier W, Sikela JM.
XX
PT      WPI: 2002-130744/17.
XX
DE      Determining human or non-human primate polynucleotide or polypeptide
PT      sequences associated with a physiological trait and have
PT      evolutionary changes, for therapeutic use, involves using
PT      methods -
XX
XX      Example 19; Fig 13; 14pp; English.
```


XX New polynucleotides encoding protease homologs of the G-protein-coupled
PT protease family, useful in identifying agonists and antagonists for
PT diagnosis and treatment of protease mediated disorders -
PS disclosure: Column 155-156; 246pp; English.

XX The invention relates to an isolated human protease nucleic acid molecule
CC comprising a nucleotide sequence of 546 base pairs, one of 266 fully
CC defined in the specification. Also disclosed are production of an
CC isolated polypeptide encoded by the nucleic acid, comprising introducing
CC the nucleic acid into a host cell and culturing under conditions to
CC express the protein from the nucleic acid, use of an antibody to
CC detect the encoded protein in a sample and to modulate its *in vivo*
CC activity, identifying agents that bind to the protein and identification
CC of a polynucleotide agent that modulates the expression of the nucleic
CC acid or its complement (i.e. gene therapy). The nucleic acid can be used
CC to identify an agent that modulates the expression or activity of the
CC nucleic acid, and can be used to isolate the protein. The nucleic acid
CC can be used in diagnostic assays for determining nucleic acid expression
CC as well as activity in the context of a biological sample (e.g., blood,
CC serum, cells, tissue) to determine whether an individual has a disease or
CC disorder, or is at risk of developing a disease or disorder, associated
CC with aberrant expression or activity of the nucleic acid. The nucleic
CC acid can be used to detect mutations in protease genes and gene
CC expression products such as mRNA. The nucleic acid can be used as
CC hybridisation probes to detect naturally-occurring genetic mutations in
CC a protease gene. The nucleic acid can be used in drug screening methods
CC to identify agonists and antagonists that can be used to diagnose and
CC treat such protease mediated disorders e.g., proliferative,
CC differentiative, developmental or hematopoietic disorders. The nucleic
CC acid can be used as probes, primers, in biological assays, to determine
CC patterns of gene expression, to design ribozymes and to construct
CC transgenic animals. The present sequence represents one of the 268
CC disclosed human G-protein-coupled protease cDNA sequences.

XX
SQ Sequence 606 BP; 129 A; 192 C; 165 G; 115 T; 5 other;

Alignment Scores:
Pred. No. 16.3 Length: 606
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
DB: 24 Gaps: 0

US-09-831-458a-12 (1-325) x ABK30341 (1-606)

QY 34 AAGCTGAGTGTGGTGGCAGCTTGGTG 225
|||||
Db 199 GCTGGAGTGTGGTGGCAGCTTGGTG 225

RESULT 30
AAT86014
ID AAT86014 standard; cDNA: 1183 BP.
AC AAT86014;
XX
DT 08-MAY-1998 (first entry)
XX
DE Human beta meltrin cDNA fragment 2.
XX
KW Meltrin: myotube formation; muscle tissue; therapy; membrane protein;
KW myoblasts; adhesion; fusion; medicinal; muscle disorder; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1183
FT /tag= a
FT /product= beta meltrin
FT /note= "partial coding sequence"
FT /codon_start= 2

XX
PN WO9731109-A1.
XX
XX 28-AUG-1997.
PD
XX 17-OCT-1996; 96WO-JP03017.
XX
XX 23-FEB-1996; 96JP-U061756.
PK
XX (MOCHI) MOCHIDA PHARM CO. LTD.
PA
XX Fujisawa A, Mizushima G, Oikawa N, Shirakawa K, Yanaka
FI
XX WPI: 1997-435161/40.
DR
XX P-IPDB: AAW5722.
XX
PT Meltrin proteins involved with myotube formation in muscle
PT metallo-proteinase and disintegrin domains
XX
PS Claim 5; FIG 17; 262pp; Japanese.
XX
CC AAT86013 and AAT86014 encode fragments of a new human, novel
CC beta meltrin, which participates in the adhesion and fusion
CC in the course of myotube formation in muscle tissue. Meltrin
CC partial peptides can be used in the study of myotube formation
CC intracellular domain, a membrane-penetrating domain, a
CC domain, a disintegrin domain, a metalloproteinase domain
CC domain and a cysteine rich domain although these are not
CC specification. Meltrin antagonists can be used as bone ad
CC inhibitors, e.g. for the treatment of osteoporosis and bone
CC and for prevention of cancer cell infiltration.

XX
SQ Sequence 1183 BP; 264 A; 419 C; 469 G; 241 T; 0 other;

Alignment Scores:
Pred. No. 40.6 Length: 1183
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
DB: 18 Gaps: 0

US-09-831-458a-12 (1-325) x AAT86014 (1-1183)

QY 34 AAGCTGAGTGTGGTGGCAGCTTGGTG 1024
|||||
Db 998 GCTGGAGTGTGGTGGCAGCTTGGTG 1024

RESULT 31
AAD03221
ID AAD03221 standard; cDNA: 1519 BP.
AC AAD03221;
XX
DT 13-JUN-2001 (first entry)
XX
DE Mouse type II integral membrane protein, cDNA.
XX
KW Mouse; type II integral membrane protein; cDNA; anti-
KW immunosuppressive; autoimmune disorder; allergy; viral in
KW antigen presenting cell; bacterial infection; macrophage;
KW DC; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 92..808
FT /tag= a
FT /product= "Mouse type II integral membrane
PN WO200119669-A1.

XX 13-MAR-1997.
 XX
 XX 05-SEP-1996; 96AW0-GR02180.
 XX
 XX 11-JUN-1996; 96GB-0012145.
 PR 05-SEP-1995; 95GB-0018026.
 PR 20-OCT-1995; 95GB-0021496.
 PR 20-OCT-1995; 95GB-0021497.
 PR 21-JUN-1995; 95GB-0026230.
 XX
 PA (CLUT) CELLTECH THERAPEUTICS LTD
 XX
 PI Dooherly AJP, Slocombe PM;
 XX
 DR WPI: 1997-192896/17.
 DR P-PSDB; AAW01825.
 XX
 PT DNA encoding human metalloproteinase - used to develop prods. for
 PT treating, e.g. infectious diseases, inflammatory disease or allergy
 XX
 PS Claim 1: Page 2-4; 17pp; English.
 XX
 CC A DNA fragment (AAT62558) includes a partial coding sequence (see also
 CC AAT62559) for a novel human metalloproteinase (AAW01825). The DNA can
 CC be used, by probing a human genomic or cDNA library with labelled
 CC oligonucleotide probes. It can be used to produce the
 CC metalloproteinase in transformed prokaryotic or eukaryotic host
 CC cells; insertion downstream of the human cytomegalovirus promoter
 CC in pEE12 plasmid vectors allows expression in CHO and NSO mouse
 CC melanoma cells. The DNA can also be used in diagnosis, e.g.
 CC to determine enzyme deficiency in a subject, or to design antisense
 CC constructs useful for the treatment of diseases or disorders
 CC related to over- or under-regulated prodn. of metalloproteinase.
 XX
 SQ Sequence 2648 BP; 620 A; 742 C; 725 G; 561 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 65.4 Length: 2648
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 DB: 18 Gaps: 0
 US-09-831-458A-12 (1-325) x AAT62558 (1-2648)
 OY 34 AAtAGtYValLeuValAlaIleLeuVal 42
 DB 1590 GCTGCACTGTCTGTCGTCATCTTGATG 1616
 RESULT 34
 AAK12899
 ID AAK12899 standard; cDNA; 2755 BP.
 XX
 AC AAK12899;
 XX
 DT 09-Apr-2002 (first entry)
 XX
 DE human protease PRTS-16 cDNA sequence.
 XX
 XX human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;
 KW cell proliferative disorder; developmental disorder; epilepsy;
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;
 KW reproductive disorder; endometriosis; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 38..2677
 FT /*tag= a

FT
 XX
 XX W0200196408-A2.
 XX
 XX 27-DEC-2001.
 PD
 XX 13-JUN-2001; 2001WO-0819178.
 PF
 XX 16-JUN-2003; 2003US-213466P.
 XX 22-JUN-2003; 2003US-213955P.
 PR 29-JUN-2003; 2003US-215466P.
 PR 07-JUL-2003; 2003US-216791P.
 PR 14-JUL-2003; 2003US-218946P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Yue H, Elliott VS, Gandhi AR, Lal P, Al-young J, T
 PI Deleageane AM, Baughn MK, Nguyen DB, Lee EA, Hatala
 PI Walla NK, Yao MG, Lu DM, Patterson C, Tang YT, Wa
 PI Azimzal Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das
 PI Kallick DA;
 XX
 DR WPI: 2003-09437/12.
 DR P-PSDB; AAU74756.
 XX
 PT Twenty one human proteases (referred to as PRTS-1 to PRTS-21) are used in the diagnosis, treatment and prevention of gastroint
 PT gastritis), cardiovascular (e.g. atherosclerosis) and c
 PT proliferative (e.g. cancer) disorders -
 XX
 PS Claim 5: Page 172-173; 177pp; English.
 XX
 CC The present invention relates to twenty one new human pr
 CC referred to as PRTS-1 to PRTS-21. The PRTS polynucleotid
 CC polypeptides of the invention are useful in the diagnos
 CC prevention of gastrointestinal e.g. gastritis, esophag
 CC Crohn's disease, cardiovascular e.g. atherosclerosis, myo
 CC myocardial infarction, autoimmune/inflammatory e.g. acq
 CC immunodeficiency syndrome (AIDS), allergies and osteoar
 CC proliferative e.g. cancer, developmental e.g. Duchenne m
 CC muscular dystrophy, epithelial e.g. dermatitis, neurolog
 CC epilepsy and Alzheimer's disease and reproductive e.g. m
 CC endometriosis disorders. Numerous other examples of eac
 CC given in the specification. The present nucleic acid se
 CC the human protease PRTS-16 protein of the invention.
 XX
 SQ Sequence 2755 BP; 633 A; 787 C; 770 G; 565 T; 0 other;
 SQ
 Alignment Scores:
 Pred. No.: 67.9 Length: 2755
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 DB: 24 Gaps: 0
 US-09-831-458A-12 (1-325) x AAK12899 (1-2755)
 OY 34 AAtAGtYValLeuValAlaIleLeuVal 42
 DB 1922 GCTGCACTGTCTGTCGTCATCTTGATG 1948
 RESULT 35
 AAC91177
 ID AAC91177 standard; DNA; 2757 BP.
 XX
 AC AAC91177;
 XX
 DT 20-MAR-2001 (first entry)
 XX
 DE ADAM gene #1.
 XX
 KW ADAM; disintegrin; metalloprotease; food additive; beta

KP ovarian; ss.
 XX
 XX Homo sapiens.
 XX
 XX W0200073423-A2.
 XX
 XX 07-DEC-2000.
 XX
 XX 25-MAY-2000; 2000MO-US14308.
 XX
 XX 27-MAY-1999; 99US-0136388.
 XX 09-JUN-1999; 99US-0142940.
 XX 28-JAN-2000; 2000US-0178717.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Wei P, Ni J, Hastings CA, Shi Y;
 XX
 XX WPI: 2001-016507/02.
 XX
 XX Seven nucleic acid molecules encoding ADAM polypeptides containing a
 XX disintegrin and metalloprotease domain, useful in the prevention,
 XX treatment and diagnosis of cancer, immune disorders, cardiovascular
 XX disorders and neurological diseases -
 XX
 XX Claim 1: Page 264-265; 287pp; English.
 XX
 XX The present invention relates to seven members of the ADAMs (proteins
 XX which contain A Disintegrin And Metalloprotease domain) protein family.
 XX The ADAMs proteins and DNA may be used to treat disease, as a food
 XX additive or preservative, for chromosome identification, as probes
 XX for diagnosing a disorder related to the female reproductive system,
 XX particularly breast and/or ovarian cancer. They are also useful in the
 XX gene therapy of breast and ovarian cancer.
 XX
 XX Sequence 2757 BP; 640 A; 803 C; 787 G; 527 T; 0 other;
 XX
 XX Alignment Scores:
 XX Pred No: 67.9 Length: 2757
 XX Score: 9.00 Matches: 9
 XX Percent Similarity: 100.00% Conservative: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 2.77% Gaps: 0
 XX DB: 22
 XX
 XX US-09-831-458A-12 (1-325) x AAC91177 (1-2757)
 XX
 XX 34 AlaclyValleuValAlaIleleVal 42
 XX ||||||||||||||||||||||||||||
 XX 2113 GCTGAGTGTGTGGTGGCCATCTTGCTG 2139
 XX
 XX R4SULT 36
 XX AAD39160
 XX ID AAD39160 standard; cDNA; 2757 BP.
 XX
 XX AC AAD39160;
 XX
 XX 04-OCT-2002 (first entry)
 XX
 XX Human novel protease cDNA #2.
 XX
 XX Human; novel human protein; NHP; protease; metelin-beta; ADAM 19;
 XX metalloproteinase; drug screening; clinical trial monitoring; cosmetic
 XX Pharmacological; nutraceutical; gene therapy; chromosome 5; gene; ss;
 XX
 XX Homo sapiens
 XX
 XX Key Location/Qualifiers
 XX CDS 1..2757
 XX F1 /*tag= a
 XX F1 /product= "Human novel protease #2"
 XX
 XX FN W0200236759-A2.

XX
 XX 10-MAY-2002.
 XX
 XX 30-OCT-2001; 2001WO-US48498.
 XX
 XX 01-NOV-2000; 2000US-244939P.
 XX
 XX (LEXI-) LEXICON GENETICS INC.
 XX
 XX Walke DW, Wilanowski NI, Friedle CJ;
 XX
 XX WPI: 2002-479761/51.
 XX F: FSDB; AAF24351.
 XX
 XX New nucleic acid encoding a protein that shares sequence similarity
 XX with mammalian metelin-beta/ADAM 19 homolog metalloprotease for use in
 XX diagnosis and treatment, and in drug screening -
 XX
 XX Disclosure, Page 17-38, 48pp, English.
 XX
 XX The invention relates to novel human proteins (NHP), novel human
 XX proteases that shares sequence similarity with mammalian metelin-beta/
 XX ADAM 19 homolog metalloprotease, and their corresponding nucleic acid
 XX sequences. NHP DNA is useful for diagnosis, drug screening, clinical
 XX trial monitoring, treatment of diseases and disorders and
 XX pharmacological, cosmetic and nutraceutical applications. It is also
 XX useful as hybridisation probes for screening libraries, and assessing
 XX gene expression patterns (particularly using a microarray or high-
 XX throughput chip format). It is also useful in restriction fragment
 XX length polymorphism (RFLP) analysis to identify specific individuals,
 XX and as a probe to identify the corresponding mutant allele in the
 XX libraries. It is also useful for the detection of mutant
 XX NHPs or inappropriately expressed NHPs for the diagnosis of disease.
 XX It is also useful in gene therapy. The present sequence is novel human
 XX protease cDNA. Human NHP gene is located at chromosome 5.
 XX
 XX Sequence 2757 BP; 641 A; 802 C; 787 G; 527 T; 0 other;
 XX
 XX Alignment Scores:
 XX Pred No: 67.9 Length: 2757
 XX Score: 9.00 Matches: 9
 XX Percent Similarity: 100.00% Conservative: 0
 XX Best Local Similarity: 100.00% Mismatches: 0
 XX Query Match: 2.77% Gaps: 0
 XX DB: 24
 XX
 XX US-09-831-458A-12 (1-325) x AAD39160 (1-2757)
 XX
 XX 34 AlaclyValleuValAlaIleleVal 42
 XX ||||||||||||||||||||||||||||
 XX 2113 GCTGAGTGTGTGGTGGCCATCTTGCTG 2139
 XX
 XX RESULT 37
 XX AAD30585
 XX ID AAD30585 standard; cDNA; 2757 BP.
 XX
 XX AC AAD30585;
 XX
 XX 21 MAY-2002 (first entry)
 XX
 XX Human protease, PKTS-18 cDNA.
 XX
 XX Human; protease; PKTS 18; enzyme; gastritis; cirrhosis; Crohn's disease;
 XX gastrointestinal disorder; autoimmune; inflammatory; cell proliferative;
 XX cardiovascular; developmental; epithelial; neurological; reproductive;
 XX AIDS; Acquired Immune Deficiency Syndrome; allergy; rheumatoid arthritis;
 XX anaemia; asthma; atherosclerosis; hypertension; myocardial infarction;
 XX hepatitis; cancer; psoriasis; Cushing's syndrome; hypothyroidism; eczema;
 XX epilepsy; Alzheimer's disease; Huntington's disease; Parkinson's disease;
 XX Pick's disease; infertility; vitiligo; drug screening; gene therapy; ss.
 XX
 XX Homo sapiens.

```

FH Key Location/Qualifiers
FT CDS 1..2757
FT /*tag= a
FT /product= "Human PRIS-18 protein"
FT stq_peptide 1..66
FT /*tag= b
FT /*tag= c
FT mat_peptide 67..2754
FT /product= "Mature PRIS 18 protein"
FN W0200298496-A2
FD 41-JAN-2002
PE 17-MAY-2001: 2001WO-0522347
XX 21-JUL-2000: 2000US-220064P
XX 28-JUL-2000: 2000US-221680P
XX 04-AUG-2000: 2000US-223544P
XX 11-AUG-2000: 2000US-224717P
XX 16-AUG-2000: 2000US-225986P
XX 23-AUG-2000: 2000US-227568P
XX (INCYTE GENOMICS INC)
XX
XX Delegrave AM, Gandhi AR, Hatalia AIA, Lu DAM, Patterson C
XX Tribouley CM, Das D, Kallick DA, Nguyen DB, Lee EA, Khan PA,
XX Yue H, Au-Yang J, Griffin JA, Policky JL, Ramkumar J, Yang J:
XX Thangavelu K, Ding L, Kearney L, Baughn MR, Borowsky ML:
XX Sanjanwalla MS, Yao MG, Burford N, Walla NK, Lal P, Lee S, Todd S:
XX Lo TP, Tang YT, Elliott VS, Azimzal Y, Lu Y:
XX WPI: 2002-206082/26.
XX P-PSDB: AAE19181.
XX
XX New human protease polypeptide, useful in diagnosis, prevention and
XX treatment of gastrointestinal, cardiovascular, autoimmune/inflammatory,
XX cell proliferative, developmental, epithelial and neurological
XX disorders -
XX
XX Claim 5: Page 179; 182pp: English.
XX
XX The invention relates to an isolated human protease polypeptide (PRIS).
XX PRIS protein and DNA are useful for diagnosing, treating and preventing
XX gastrointestinal disorders (gastritis, cirrhosis, Crohn's disease),
XX autoimmune/inflammatory disorders (AIDS, allergy, rheumatoid arthritis,
XX anemia, asthma), cardiovascular disorder (atherosclerosis, hypertension,
XX myocardial infarction), cell proliferative disorders (hepatitis, cancer,
XX psoriasis), developmental disorders (Cushing's syndrome, hypothyroidism),
XX epithelial disorder (vitiligo, keloid, eczema), neurological disorders
XX (epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease,
XX Parkinson's disease), and reproductive disorders (infertility). PRIS
XX protein is useful in a number of drug screening techniques and to
XX analyse the proteome of a tissue or cell type. PRIS DNA is useful for
XX creating knockin humanised animals or transgenic animals to model human
XX diseases, in somatic or germline gene therapy and in microarrays
XX utilising fluids or tissues from patients to detect altered PKIN
XX expression. The present sequence is human PRIS-18 cDNA.
XX
XX Sequence 2757 BP; 641 A; 802 C; 787 G; 527 T; 0 other.
XX
XX Alignment Scores:
XX Pred. No.: 67.9
XX Score: 9.00
XX Percent Similarity: 100.00%
XX Best Local Similarity: 100.00%
XX Query Match: 2.77%
XX DB: 24
XX Gaps: 0
XX
XX US-09-831-458A-12 (1-325) x AAD30585 (1-2757)
XX
XX 34 AlaglyValLeuValAlaIleLeuVal 42
XXXXXXXXXXXXXXXXXXXXXXXXXXXX

```

```

DB 2113 GCTGCAATGTTGCTGGCCATCTTATG 2149
RESULT 38
ID AAD39159
ID AAD39159 standard; cDNA: 2781 BP.
XX AAD39159:
XX
XX 04-OCT-2002 (first entry)
XX
XX Human novel protease cDNA #1.
XX
XX Human novel human protein; NHP: protease; meltrin-beta;
XX metalloprotease; drug screening; clinical trial monitor in
XX pharmacological; multicentric; gene therapy; chromosome-
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..2781
XX /*tag= a
XX /product= "Human novel protease #1"
XX
XX W0200236759-A2.
XX
XX 10-MAY-2002.
XX
XX 40-OCT-2001: 2001WO-054849B.
XX
XX 01-NOV-2000: 2000US-244939P.
XX
XX (LEXI- ) LEXICON GENETICS INC.
XX
XX Walke DW, Wilgowski NL, Frittle CJ:
XX WPI: 2002-474761/51.
XX P-PSDB: AAE24350.
XX
XX New nucleic acid encoding a protein that shares sequence
XX with mammalian meltrin-beta/ADAM 19 homolog metalloprote
XX diagnosis and treatment, and in drug screening -
XX
XX Disclosure, Page 34-35; 48pp: English.
XX
XX The invention relates to novel human proteins (NHP), novel
XX proteases that shares sequence similarity with mammalian
XX ADAM 19 homologue metalloprotease, and their correspond
XX sequences. NHP DNA is useful for diagnosis, drug screen
XX trial monitoring, treatment of diseases and disorders in
XX pharmacological, cosmetic and nutraceutical applications.
XX useful as hybridisation probes for screening libraries.
XX gene expression patterns (particularly using a microarray
XX throughput chip format). It is also useful in restriction
XX length polymorphism (RFLP) analysis to identify specific
XX and as a probe to identify the corresponding mutant alle
XX libraries. It is also useful for the detection of mutat
XX NHPs or inappropriately expressed NHPs for the diagnosis
XX it is also useful in gene therapy. The present sequence
XX protease cDNA. Human NHP gene is located at chromosome
XX
XX Sequence 2781 BP; 644 A; 807 C; 794 G; 534 T; 0 other:
XX
XX Alignment Scores:
XX Pred. No.: 68.5
XX Score: 9.00
XX Percent Similarity: 100.00%
XX Best Local Similarity: 100.00%
XX Query Match: 2.77%
XX DB: 24
XX Gaps: 0
XX
XX US-09-831-458A-12 (1-325) x AAD39159 (1-2781)
XX
XX 34 AlaglyValLeuValAlaIleLeuVal 42
XXXXXXXXXXXXXXXXXXXXXXXXXXXX

```

```

|||||
Dn 2113 GCTGGAGTGTGTGGCAGCTTGTG 2139
RESULT 39
A039162
1) AAD39162 standard; cDNA; 2868 BP.
XX
AC AAD39162:
XX
DE 04-OCT-2002 (first entry)
XX
DE Human novel protease cDNA #4.
XX
KM Human; novel human protein; NHP; protease; meltrin-beta; ADAM 19;
KV metalloprotease; drug screening; clinical trial monitoring; cosmetic;
KW pharmacological; nutraceutical; gene therapy; chromosome 5; gene; SS.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2868
FT /tag= a
FT /product= "human novel protease #4"
XX
DE 04-OCT-2001; 2001WO-US48498.
XX
DE 10-MAY-2002.
XX
DE 01-NOV-2000; 2000US-244939P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Wilganowski NL, Fridde CJ:
XX
DE WPI: 2002-479761/51.
XX
DE P-PSDB: AAE24353.
XX
PI New nucleic acid encoding a protein that shares sequence similarity
PI with mammalian meltrin-beta/ADAM 19 homolog metalloprotease for use in
PI diagnosis and treatment, and in drug screening -
XX
PS Disclosure: Page 44; 48pp; English.
XX
CC The invention relates to novel human proteins (NHP), novel human
CC proteases that shares sequence similarity with mammalian meltrin-beta/
CC ADAM 19 homolog metalloprotease, and their corresponding nucleic acid
CC sequences. NHP DNA is useful for diagnosis, drug screening, clinical
CC trial monitoring, treatment of diseases and disorders and
CC pharmacological, cosmetic and nutraceutical applications. It is also
CC useful as hybridisation probes for screening libraries, and assessing
CC gene expression patterns (particularly using a microarray or high-
CC throughput chip format). It is also useful in restriction fragment
CC length polymorphism (RFLP) analysis to identify specific individuals,
CC and as a probe to identify the corresponding mutant allele in the
CC libraries. It is also useful for the detection of mutant
CC NHPs or inappropriately expressed NHPs for the diagnosis of disease.
CC It is also useful in gene therapy. The present sequence is novel human
CC protease cDNA. Human NHP gene is located at chromosome 5.
XX
XX Sequence 2868 BP; 668 A; 826 C; 821 G; 550 T; 0 other:
XX
Alignment Scores:
Pred. No.: 70.5 Length: 2868
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
DB: 24 Gaps: 0
US-09-831-458A-12 (1-325) x AAD39162 (1-2868)

```

```

OY 34 AbaglyValienValiallslanVal 42
|||||
Dn 2113 GCTGGAGTGTGTGGCAGCTTGTG 2139
RESULT 40
A039161
ID AAD39161 standard; cDNA; 2892 BP.
XX
AC AAD39161:
XX
DE 04-OCT-2002 (first entry)
XX
DE Human novel protease cDNA #3.
XX
KM Human; novel human protein; NHP; protease; meltrin-beta; ADAM 19;
KV metalloprotease; drug screening; clinical trial monitoring; cosmetic;
KW pharmacological; nutraceutical; gene therapy; chromosome 5; gene; SS.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2892
FT /tag= a
FT /product= "human novel protease #3"
XX
DE 04-OCT-2001; 2001WO-US48498.
XX
DE 10-MAY-2002.
XX
DE 01-NOV-2000; 2000US-244939P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Walke DW, Wilganowski NL, Fridde CJ:
XX
DE WPI: 2002-479761/51.
XX
DE P-PSDB: AAE24352.
XX
PI New nucleic acid encoding a protein that shares sequence similarity
PI with mammalian meltrin-beta/ADAM 19 homolog metalloprotease for use in
PI diagnosis and treatment, and in drug screening -
XX
PS Claim 2; Page 40-41; 48pp; English.
XX
CC The invention relates to novel human proteins (NHP), novel human
CC proteases that shares sequence similarity with mammalian meltrin-beta/
CC ADAM 19 homolog metalloprotease, and their corresponding nucleic acid
CC sequences. NHP DNA is useful for diagnosis, drug screening, clinical
CC trial monitoring, treatment of diseases and disorders and
CC pharmacological, cosmetic and nutraceutical applications. It is also
CC useful as hybridisation probes for screening libraries, and assessing
CC gene expression patterns (particularly using a microarray or high-
CC throughput chip format). It is also useful in restriction fragment
CC length polymorphism (RFLP) analysis to identify specific individuals,
CC and as a probe to identify the corresponding mutant allele in the
CC libraries. It is also useful for the detection of mutant
CC NHPs or inappropriately expressed NHPs for the diagnosis of disease.
CC It is also useful in gene therapy. The present sequence is novel human
CC protease cDNA. Human NHP gene is located at chromosome 5.
XX
XX Sequence 2892 BP; 671 A; 834 C; 828 G; 559 T; 0 other:
XX
Alignment Scores:
Pred. No.: 71.1 Length: 2892
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.77% Indels: 0
DB: 24 Gaps: 0
US-09-831-458A-12 (1-325) x AAD39161 (1-2892)

```


PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232091.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0234447.
 PR 14-SEP-2000; 2000US-0233398.
 PR 14-SEP-2000; 2000US-0233399.
 PR 14-SEP-2000; 2000US-0234400.
 PR 14-SEP-2000; 2000US-0234401.
 PR 14-SEP-2000; 2000US-0234064.
 PR 14-SEP-2000; 2000US-0234064.
 PR 14-SEP-2000; 2000US-0234064.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234967.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0234984.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236357.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236464.
 PR 29-SEP-2000; 2000US-0236470.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237049.
 PR 13-OCT-2000; 2000US-0239035.
 PR 13-OCT-2000; 2000US-0239037.
 PR 20-OCT-2000; 2000US-0240090.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241898.
 PR 20-OCT-2000; 2000US-0241899.
 PR 20-OCT-2000; 2000US-0241899.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246612.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251040.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251956.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR (HUMAN) HUMAN GENOME SCI (HG).
 PR Person (A), Parish (S), Eubank (SM).
 PR WPI: 2001-45566/50.
 PR Re of polypeptides and polypeptides useful for diagnosing,
 PR preventing, treating (neural, immune system, muscular, reproductive,
 PR pulmonary, cardiovascular, renal, proliferative disorders and cancerous
 PR diseases.
 PR
 PR Discovered, SEQ ID NO 2285, 1180pp, English.
 PR
 PR The present invention relates to the isolation of novel human enzyme
 PR polypeptides (AA022915-AA023014), and the cDNA and genomic sequences
 PR encoding them. The enzyme polypeptides or the invention may comprise the
 PR functional classes of oxidoreductases, transferases, hydrolases, lyases,
 PR isomerases or ligases. The sequences of the invention are useful in the
 PR diagnosis, treatment, prevention and/or prognosis of a wide range of
 PR disorders including hyperproliferative disorders (e.g. cancer),
 PR immunodeficiency disorders (e.g. AIDS), autoimmune disorders
 PR (e.g. arthritis), neurological disorders (e.g. Alzheimer's disease),
 PR metabolic disorders (e.g. phenylketonuria), inflammatory disorders
 PR (e.g. psoriasis), cardiovascular disorders (e.g. atherosclerosis),
 PR blood related disorders (e.g. haemophilia), reproductive disorders
 PR (e.g. infertility) and infectious disorders (e.g. influenza). The
 PR polypeptides of the invention can also be used in gene therapy.
 PR AA021697-AA021892 represent DNA sequences encoding for the novel human
 PR enzyme polypeptides of the invention.
 PR Note: The sequence data for this patent did not form part of the printed
 PR specification, but was obtained in electronic format directly from WIPO
 PR at http://seq.ip/seq/pub/published_pat_sequences.
 PR
 PR Seq-Info: 119 bp, 1007 A, 742 C, 814 G, 1475 T, 0 other.
 PR
 PR Alignment Scores:
 PR Pred. No.: 105 Length: 4418
 PR Score: 7.00 Matches: 9
 PR Percent Similarity: 100.00% Conserved: 0
 PR Best Local Similarity: 100.00% Mismatches: 0
 PR Query Match: 7.77% Indels: 0
 PR DB: 22 Gaps: 0
 PR
 PR US 09-831 458A 12 (1 225) x AA02169 (1-4418)
 PR
 PR 27 47 VVHGGSTGRIEUSGRIHGLDGLN 55
 PR 28 4382 GTPGGMIAAGCTGTGGVAGAAVGA 4356
 PR
 PR RESULT 43
 PR AAK68448
 PR ID AAK69148 standard, DNA, 16424 BP.
 PR XX AAK68448;
 PR AC AAK68448;
 PR XX
 PR DT 06 NOV 2001 (first entry)
 PR XX


```

XX  (HOMA- ) HUMAN GENOME SCI INC.
XX  Rosen CA, Barash SC, Ruben SM,
XX
XX  WPI: 2001-483426/52
XX
XX  Nucleic acids encoding human immune/hematopoietic antigen p31/p31bides,
XX  useful for preventing, diagnosing and/or treating cancers and
XX  metastasis -
XX
XX  Disclosure: SEQ ID NO 23260; 367bp - Sequence Listing; English.
XX
XX  AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (1)
XX  amino acid sequences given in AAK62176 to AAK61921 (1) have cytostatic
XX  activity, and can be used in gene therapy and vaccine production; (1)
XX  proteins and polynucleotides may be used in the prevention, diagnosis and
XX  treatment of diseases associated with inappropriate (1) expression. For
XX  example, they may be used to treat disorders associated with decreased
XX  expression by rectifying mutations or deletions in a patient's genome
XX  that affect the activity of (1) by expressing inactive proteins or to
XX  supplement the patients' own production of (1). Additionally, (1)
XX  polynucleotides may be used to produce the secreted (1), by inserting
XX  the nucleic acids into a host cell and culturing the cell to express the
XX  protein. (1) proteins and polynucleotides may be used to prevent, the
XX  disease and treat immune/hematopoietic-related diseases, especially
XX  cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX  to AAK6764 represent human immune/hematopoietic antigen genomic
XX  sequences from the present invention. AAK54942 to AAK54950 and AAK62169
XX  represent sequences used in the exemplification of the present invention.
XX
XX  Sequence 16424 BP: 4245 A; 3614 G; 4136 C; 4469 T; 0 other.
XX
XX  Alignment Scores:
XX  Prod. No.: 366
XX  Percent Similarity: 9.00 Length: 16424
XX  Percent Similarity: 100.00% Conservative: 9
XX  Best Local Similarity: 100.00% Mismatches: 0
XX  Query Match: 2.77% Indels: 0
XX  Dbs: 22 Gaps: 0
XX
XX  US 09-831-456A 12 (1 325) X AAK58446 (1 16424)
XX
XX  34 A1ac3yVal16uVal1a1a1a1eLeval 42
XX  ||||||||||||||||||||||||||||
XX  Db 12545 GCTGCACTGTCGTCGTCGCACTGCTGCT 12571
XX
XX  RESULT 44
XX  AAK75677/
XX  ID AAK75677 standard; DNA: 16424 BP.
XX
XX  AAK75677:
XX  07-NOV-2001 ((first entry))
XX
XX  Human immune/hematopoietic antigen genomic sequence SEQ ID NO:23260.
XX
XX  Human: immune; hematopoietic; immune/hematopoietic antigen; cancer;
XX  cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX  Homo sapiens.
XX
XX  W0200157182-A2.
XX
XX  09-AUG-2001.
XX
XX  17 JAN 2001: 2601W-050154
XX
XX  31-JAN-2000: 2600WS-017005
XX  04-FEB-2000: 2600WS-018002

```

FR	24-SEP-2000	2000005-0184664
FR	02-MAR-2000	2000005-0186300
FR	16-MAR-2000	2000005-0186874
FR	11-MAR-2000	2000005-0190076
FR	18-APR-2000	2000005-0198113
FR	19-MAY-2000	2000005-0200515
FR	37-JUN-2000	2000005-0205417
FR	28-JUN-2000	2000005-0214886
FR	06-JUN-2000	2000005-0215135
FR	07-JUL-2000	2000005-0216647
FR	07-JUL-2000	2000005-0217680
FR	11-JUL-2000	2000005-0217449
FR	11-JUL-2000	2000005-0218290
FR	14-JUL-2000	2000005-0219064
FR	26-JUL-2000	2000005-0220964
FR	14-AUG-2000	2000005-0224519
FR	14-AUG-2000	2000005-0224518
FR	14-AUG-2000	2000005-0225214
FR	14-AUG-2000	2000005-0225266
FR	14-AUG-2000	2000005-0225367
FR	14-AUG-2000	2000005-0225368
FR	14-AUG-2000	2000005-0225270
FR	14-AUG-2000	2000005-0225447
FR	14-AUG-2000	2000005-0225757
FR	14-AUG-2000	2000005-0225758
FR	14-AUG-2000	2000005-0225759
FR	18-AUG-2000	2000005-0226279
FR	22-AUG-2000	2000005-0226641
FR	22-AUG-2000	2000005-0226868
FR	22-AUG-2000	2000005-0227212
FR	22-AUG-2000	2000005-0227219
FR	22-AUG-2000	2000005-0228924
FR	30-AUG-2000	2000005-0229287
FR	01-SEP-2000	2000005-0229343
FR	01-SEP-2000	2000005-0229343
FR	01-SEP-2000	2000005-0229345
FR	05-SEP-2000	2000005-0229509
FR	05-SEP-2000	2000005-0229513
FR	06-SEP-2000	2000005-0230347
FR	06-SEP-2000	2000005-0230348
FR	08-SEP-2000	2000005-0231242
FR	08-SEP-2000	2000005-0231243
FR	08-SEP-2000	2000005-0231244
FR	08-SEP-2000	2000005-0231244
FR	14-SEP-2000	2000005-0232398
FR	14-SEP-2000	2000005-0232399
FR	14-SEP-2000	2000005-0232400
FR	08-SEP-2000	2000005-0231414
FR	08-SEP-2000	2000005-0231080
FR	08-SEP-2000	2000005-0232081
FR	14-SEP-2000	2000005-0232401
FR	14-SEP-2000	2000005-0233064
FR	14-SEP-2000	2000005-0233065
FR	21-SEP-2000	2000005-0234253
FR	21-SEP-2000	2000005-0234254
FR	21-SEP-2000	2000005-0234257
FR	25-SEP-2000	2000005-0234997
FR	25-SEP-2000	2000005-0234998
FR	26-SEP-2000	2000005-0235484
FR	26-SEP-2000	2000005-0235484
FR	27-SEP-2000	2000005-0235836
FR	29-SEP-2000	2000005-0236357
FR	29-SEP-2000	2000005-0236357
FR	29-SEP-2000	2000005-0236358
FR	29-SEP-2000	2000005-0236359
FR	29-SEP-2000	2000005-0236370
FR	02-OCT-2000	2000005-0236802
FR	02-OCT-2000	2000005-0237037
FR	02-OCT-2000	2000005-0237038

PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237049.
 PR 13-OCT-2000; 2000US-0239315.
 PR 13-OCT-2000; 2000US-0239347.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241231.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251889.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX PA
 XX PI Rosen CA, Farash SC, Ruben SM;
 XX WPI: 2001-483426/52.
 XX DR
 XX PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 XX usefull for preventing, diagnosing and/or treating cancers and
 XX metastasis.
 PS Disclosure; SEQ ID NO 30489; 3071pp + Sequence Listing; English.

XX AAK54941 to AAK64702 encode the human immune/hematopoietic
 CC amino acid sequences given in AAK64702 to AAK64921. (1) For
 CC activity, and can be used in gene therapy and vaccine preparation.
 CC proteins and polynucleotides may be used in the prevention and
 CC treatment of diseases associated with inappropriate (1) For
 CC example, they may be used to treat disorders associated with
 CC expression by rectifying mutations or deletions in a particular
 CC that affect the activity of (1) by expressing inactive protein
 CC supplement the patients own production of (1). Additionally,
 CC polynucleotides may be used to produce the secreted (1) in
 CC the nucleic acids into a host cell and culturing the cell.
 CC protein. (1) proteins and polynucleotides may be used to
 CC diagnose and treat immune/hematopoietic related diseases
 CC cancers and cancer metastases of hematopoietic-derived cells
 CC to AAK6794 represent human immune/hematopoietic antigen
 CC sequences from the present invention. AAK54942 to AAK5494
 CC represent sequences used in the exemplification of the present
 XX sequence 10424 bp; 4399 A; 4136 C; 3644 G; 4245 T; 0 with
 SV
 Alignment Scores:
 Pled. No.: 366 Length: 10424
 Score: 9.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.77% Indels: 0
 Dels: 22 Gaps: 0
 US-09-831-458A-12 (1-325) X AAK75677 (1-16124)
 CY 34 AAGlyValLeuValAlaIleLeuVal 42
 DB 3886 G-TGACGATGATGATGATGATGATGATG 3854
 RESULT 45
 AAS98600/C
 LU AAS98600 standard; DNA: 140036 bp.
 XX AAS98600;
 AC
 XX 12-MAR-2002 (first entry)
 DE Human genomic DNA for PHLP/NLRP.
 XX
 KW PHLP: Pleckstrin homology domain-interacting protein; NMR
 KW neuronal differentiation-related protein; insulin receptor
 KW IRS; signal transducer and activator of transcription; S1
 KW transgenic animal; diabetes mellitus type 2; hyperglycemia
 KW myotonic muscular dystrophy; acanthosis nigricans; retinopathy
 KW nephropathy; arteriosclerosis; peripheral arterial disease;
 KW adrenocarcinoma; leukemia; breast cancer; prostate cancer;
 KW ovarian cancer; autoimmune disease; inflammation; immunosuppression.
 XX Homo sapiens.
 OS
 XX WU00085785-A2.
 EN 15-NOV-2001
 EU 10-MAY-2001; 2001WO-CA00673.
 XX 11-MAY-2000; 2000US-203561P.
 XX (ROZAKIS) ROZAKIS-ADCOCK M.
 PA (FARHANG) FARHANG-FALLAH J.
 PA (CHENG) CHENG A.
 XX ROZAKIS-Adcock M, Farhang-Fallah J, Cheng A;
 PI WPI: 2002-041586/05.
 XX Novel Pleckstrin homology domain interacting protein rec

PI proteins of insulin receptor substrate family, and signal transducer
PI and activator of transcription factors to their receptors, useful to
PI treat diabetes

PS Disclosure: Page 99-133; 139pp; English.

XX The invention relates to an isolated pleckstrin homology domain
CC interacting protein (PHIP) that recruits proteins of the insulin
CC receptor substrate (IRS) family, and signal transducer and activator of
CC transcription (STAT) transcription factors to receptors that interact
CC with and phosphorylate the proteins and STAT transcription factors,
CC the nucleic acid encoding PHIP (nPHIP), a nucleic acid which binds to
CC nPHIP or regions of it, analogues, fragments or allelic variants of PHIP
CC or nPHIP, a nucleic acid sequence having substantial sequence identity or
CC sequence similarity with a nucleic acid sequence fully defined
CC human neuronal differentiation-related protein (NURF) nucleic acid
CC sequence or its exons as given in the specification, expression
CC vectors and host cells expressing the nucleic acids, anti-PHIP antibodies,
CC and a transgenic animal not already expressing PHIP. The nucleic
CC acids, proteins and antibodies are useful for diagnosis and treatment of
CC a condition associated with an insulin receptor (e.g. diabetes mellitus
CC type 2, hyperglycaemia, myotonic muscular dystrophy, acanthosis,
CC nigricans, retinopathy, nephropathy, arteriosclerosis, peripheral
CC arterial disease) or cancer (e.g. adenocarcinoma, leukaemia, breast
CC cancer, prostate cancer, colon cancer, ovarian cancer and many others
CC given in the specification), autoimmune disease, inflammation and
CC immunodeficiency. The protein is also useful for discovering
CC or testing compounds which may be either enhancers or inhibitors of PHIP
CC function. The present sequence is genomic DNA encoding PHIP and NURF.
XX
SO Sequence 140036 RF, 48506 A, 25843 C, 22459 G, 43134 T, 0 other;

Alignment Scores:

Pred. No.:	2	77e+03	Length:	140036
Score:	9.00		Matches:	9
Percent Similarity:	100.00%		Conservative:	0
Best Local Similarity:	100.00%		Mismatches:	0
Query Match:	2.77%		Indels:	0
DB:	24		Gaps:	0

US-09-831-458a-12 (1-345) X AAS9800V (1-140036)

QY 74 G1uLeuSerGluLysSerLysLeuGln 82

Db 107457 GAGTAAAGTGAAGAAATCTAACTACAA 107431

Search completed: December 7, 2002, 13:53:17
Job time : 353 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM protein - nucleic search, using frame_plus_F2u model

Run on: December 7, 2002, 11:31:17 Search time: 55 seconds
(without alignments)

1812.181 Million cell updates/sec

Title: US-09-831-458a-12

Sequence: 1 MSUSKEPRVQGLIGLCTGH LCTVINWLVKKPAAPM E 325

Scoring table:

OLIGO	60.0	Xgapext 60.0
Xgapop 60.0	Xgapext 60.0	
fgapop 6.0	fgapext 7.0	
delop 6.0	delext 7.0	

Searched: 44162 seqs, 15338381 residues

Word size: 5

Total number of hits satisfying chosen parameters: 32521

Minimum db seq length: 0

Maximum db seq length: 20000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_F2u.model -DEV=xlh
-Q=/cgm2_1/ospfn.spool/US09831458/runat_05122002_103417_16789/3pf_query.fasta_1.519
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=oli.rmi -MINMATCH=0 -1 -LOCAL=1
-LOCALEXT=0 -ONITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR_SCORE=quality -THR_MIN=5 -ALIGN=45 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=20000000
-USPR=US09831458.acfm_1.1.17 -runat_05122002_103417_16789 -NCPU 6 -TPR=3
-NO_XLPRY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMECUT=120
-WARN_TIMECUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YCAPOP=60 -YCAPEXT=60 -DELCP=6 -DELEXT=7

Database: Issued_Patents_NA:*

1: /cgm2_6/prodata/1/ina/5A.COMB.seq:*

2: /cgm2_6/prodata/1/ina/5B.COMB.seq:*

3: /cgm2_6/prodata/1/ina/5A.COMB.seq:*

4: /cgm2_6/prodata/1/ina/5A.COMB.seq:*

5: /cgm2_6/prodata/1/ina/5A.COMB.seq:*

6: /cgm2_6/prodata/1/ina/5A.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match length	DB ID	Description
1	168	51.7	1643	US-09-517-605-16
2	78	24.0	1212	US-09-591-435-3
3	78	24.0	1312	US-09-517-605-1
4	43	13.2	1212	US-09-591-435-10
5	35	10.8	1212	US-09-591-435-11
6	9	2.8	606	US-09-280-116-11
7	9	2.8	2251	US-08-836-443-2
8	9	2.8	2648	US-08-836-443-1
9	8	2.5	3744	US-08-348-353-16
10	8	2.5	3744	US-08-348-353-16
11	8	2.5	3744	US-08-465-965-16
12	8	2.5	8075	US-08-465-966-16
				US-08-374-077C-1
				Sequence 1, Appli

13	8	2.5	8075	US-08-835-690-1	Appl
14	8	2.5	8075	US-09-539-879A-1	Appl
15	7	2.2	132	US-08-450-945-68	Appl
16	7	2.2	132	US-08-476-161-68	Appl
17	7	2.2	364	US-07-744-518-24	Appl
18	7	2.2	407	US-09-280-116-105	Appl
19	7	2.2	452	US-07-662-198B-1	Appl
20	7	2.2	452	US-08-342-742-1	Appl
21	7	2.2	462	US-08-450-945-65	Appl
22	7	2.2	462	US-08-476-161-66	Appl
23	7	2.2	469	US-08-458-084-5	Appl
24	7	2.2	469	US-08-458-084-5	Appl
25	7	2.2	469	US-08-458-084-5	Appl
26	7	2.2	469	US-08-458-084-5	Appl
27	7	2.2	469	US-08-458-084-5	Appl
28	7	2.2	538	US-08-458-084-5	Appl
29	7	2.2	538	US-08-458-084-5	Appl
30	7	2.2	538	US-08-458-084-5	Appl
31	7	2.2	612	US-07-205-169-1	Appl
32	7	2.2	679	US-08-458-084-5	Appl
33	7	2.2	679	US-08-458-084-5	Appl
34	7	2.2	679	US-08-458-084-5	Appl
35	7	2.2	679	US-08-458-084-5	Appl
36	7	2.2	690	US-08-458-084-5	Appl
37	7	2.2	690	US-08-458-084-5	Appl
38	7	2.2	763	US-08-458-084-5	Appl
39	7	2.2	763	US-08-458-084-5	Appl
40	7	2.2	763	US-08-458-084-5	Appl
41	7	2.2	763	US-08-458-084-5	Appl
42	7	2.2	763	US-08-458-084-5	Appl
43	7	2.2	791	US-09-051-596-5	Appl
44	7	2.2	936	US-09-416-509C-2	Appl
45	7	2.2	982	US-08-973-068-28	Appl

ALIGNMENTS

RESULT 1
US-09-517-605-16
Sequence 16, Application: US-09-517-605-16
Patent No. 6391567
GENERAL INFORMATION:
APPLICANT: Liltman, Dan R.
APPLICANT: Kwon, Douglas S.
APPLICANT: van Kooyk, Yvette
APPLICANT: Geljendek, Theo
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RE:
FILE REFERENCE: 1049-1-017
CURRENT APPLICATION NUMBER: US-09-517-605
CURRENT FILING DATE: 2000-07-02
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
Seq ID No 16
LENGTH: 1643
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (25)..(915)
US-09-517-605-16

Alignment Scores:
Pred. No.: 3.24e-160
Score: 168.00
Percent Similarity: 100.00
Best Local Similarity: 100.00
Query Match: 51.698
DB: 4
US-09-831-458a-12 (1-325) x US-09-517-605-16 (1-1643)
Oy 1 MetSerAspSerIlySGIUpProVaValcInuInleuGlyLeuLeuGlyGly

```

1b 25 ATAGAGTACTTCCTAGAGAAATTAAGGCTGATGATGAGGCTTCTGAGGCTATAT 64
    |||
0y 21 CTATATATATATATATATATATATATATATATATATATATATATATATATAT 40
    |||
1b 85 GGGGCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 144
    |||
0y 41 TATATATATATATATATATATATATATATATATATATATATATATATATAT 60
    |||
1b 145 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 204
    |||
0y 61 TATATATATATATATATATATATATATATATATATATATATATATATATAT 80
    |||
1b 205 ATCTACACAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 264
    |||
0y 81 TATATATATATATATATATATATATATATATATATATATATATATATATAT 100
    |||
1b 265 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 324
    |||
0y 101 TATATATATATATATATATATATATATATATATATATATATATATATATAT 120
    |||
1b 325 AAATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384
    |||
0y 121 TATATATATATATATATATATATATATATATATATATATATATATATATAT 140
    |||
1b 385 TATATATATATATATATATATATATATATATATATATATATATATATATAT 444
    |||
0y 141 TATATATATATATATATATATATATATATATATATATATATATATATATAT 160
    |||
1b 445 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504
    |||
0y 161 TATATATATATATATATATATATATATATATATATATATATATATATATAT 180
    |||
1b 505 AAATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 528
    |||

```

RESULT 2

```

US-09-591-435-9
: Sequence 9, Application US/09591435
: Patent No. 6280953
: GENERAL INFORMATION:
: APPLICANT: MESSIER, WALTER
: APPLICANT: SIKELA, JAMES M
: TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
: TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
: FILE REFERENCE: GENO 200.2
: CURRENT APPLICATION NUMBER: US/09/591.435
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/591.435
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/240.915
: PRIOR FILING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: 60/073.263
: PRIOR FILING DATE: 1998-01-30
: PRIOR APPLICATION NUMBER: 60/098.987
: PRIOR FILING DATE: 1998-04-02
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 9
: LENGTH: 1212
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-591-435-9

```

Alignment Scores:

```

Pred. No. 1 380-69 Length: 1212
Score: 78.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-831-458a-12 (1-325) x US-09-591-435-9 (1 1212)

```

0y 58 GTATATATATATATATATATATATATATATATATATATATATATATATATAT 77
    |||
1b 226 TATATATATATATATATATATATATATATATATATATATATATATATATAT 279
    |||
0y 78 TATATATATATATATATATATATATATATATATATATATATATATATATAT 97
    |||
1b 280 AAATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 339
    |||
0y 98 TATATATATATATATATATATATATATATATATATATATATATATATATAT 117
    |||
1b 440 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 399
    |||
0y 118 TATATATATATATATATATATATATATATATATATATATATATATATATAT 135
    |||
1b 440 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 453
    |||

```

RESULT 3

```

US-09-517-605-1
: Sequence 1, Application US/09517605
: Patent No. 6391567
: GENERAL INFORMATION:
: APPLICANT: Littman, Dan R.
: APPLICANT: Kwan, Douglas S.
: APPLICANT: van Kooyk, Yvette
: APPLICANT: Gelfondbeck, Theo
: TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
: TITLE OF INVENTION: CELLS
: FILE REFERENCE: 1049-1-017
: CURRENT APPLICATION NUMBER: US/09/517.605
: NUMBER OF SEQ ID NOS: 17
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 1312
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (42)..(1253)
US-09-517-605-1

```

```

Alignment Scores:
Pred. No. 1 480-69 Length: 1312
Score: 78.00 Matches: 78
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 24.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-831-458a-12 (1-325) x US-09-517-605-1 (1 1312)

```

0y 58 GTATATATATATATATATATATATATATATATATATATATATATATATATAT 77
    |||
1b 261 TATATATATATATATATATATATATATATATATATATATATATATATATAT 329
    |||
0y 78 TATATATATATATATATATATATATATATATATATATATATATATATATAT 97
    |||
1b 321 AAATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 380
    |||
0y 98 TATATATATATATATATATATATATATATATATATATATATATATATATAT 117
    |||
1b 381 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 440
    |||

```

```

0y 118 TATATATATATATATATATATATATATATATATATATATATATATATATAT 135
    |||
1b 441 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 494
    |||

```

RESULT 4

```

US-09-591-435-10
: Sequence 10, Application US/09591435
: Patent No. 6280953
: GENERAL INFORMATION:

```

```

: APPLICANT: MESSIER, WALTER
: APPLICANT: SIKELA, JAMES M
: TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
: TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
: TITLE OF INVENTION: AND MEDICAL CONDITIONS
: FILE REFERENCE: CNO 200 2
: CURRENT APPLICATION NUMBER: US/09/591,435
: CURRENT FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/591,435
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/240,915
: PRIOR FILING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: 60/073,263
: PRIOR FILING DATE: 1998-01-30
: PRIOR APPLICATION NUMBER: 60/098,987
: PRIOR FILING DATE: 1998-09-02
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 10
: LENGTH: 1212
: TYPE: DNA
: ORGANISM: Pan troglodytes
: US-09-591-435-10

Alignment Scores:
Pred. No.: 2,7e-34      Length: 1212
Score: 43.00           Matches: 43
Percent Similarity: 100.00%   Conservat: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 13.23%           Indels: 0
DB: 4                  Gaps: 0

US-09-831-458a-12 (1-325) x US-09-591-435-10 (1-1212)

QY 61 11etYrGlnAsnLeuThrGlnLeuLysAlaAlaValGlyGluLeuSerGluLysSerLys 80
: |||
DB 229 AITACCAAGCAACGACCCACCTTAAGGCTGAGCTGCTGAGGAATCCAG 288

QY 81 LeuGlnGluLeuTyrosGlnLeuThrGlnLeuLysAlaAlaValGlyGluLeuProGlu 100
: |||
DB 289 CTGTAGAGAGATTAAGTAAAGATTAACCTTAACCTTAACCTTAACCTTAACCTTAAC 342

QY 101 LysSerLys 103
: |||
DB 349 AATATCTAAG 357

RESULT 5
: US-09-591-435-11
: Sequence 11, Application US/09591435
: Patent No. 6280953
: GENERAL INFORMATION:
: APPLICANT: MESSIER, WALTER
: APPLICANT: SIKELA, JAMES M
: TITLE OF INVENTION: METHODS TO IDENTIFY POLYNUCLEOTIDE AND POLYPEPTIDE
: TITLE OF INVENTION: SEQUENCES WHICH MAY BE ASSOCIATED WITH PHYSIOLOGICAL
: FILE REFERENCE: GENO.200.2
: CURRENT APPLICATION NUMBER: US/09/591,435
: CURRENT FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/591,435
: PRIOR FILING DATE: 2000-06-09
: PRIOR APPLICATION NUMBER: 09/240,915
: PRIOR FILING DATE: 1999-01-29
: PRIOR APPLICATION NUMBER: 60/073,263
: PRIOR FILING DATE: 1998-01-30
: PRIOR APPLICATION NUMBER: 60/098,987
: PRIOR FILING DATE: 1998-09-02
: NUMBER OF SEQ ID NOS: 13
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 11
: LENGTH: 1212
: TYPE: DNA
: ORGANISM: Gorilla gorilla

```

```

US-09-591-435-11
Alignment Scores:
Pred. No.: 3,15e-26      Length: 1212
Score: 35.00           Matches: 35
Percent Similarity: 100.00%   Conservat: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 10.77%           Indels: 0
DB: 4                  Gaps: 0

US-09-831-458a-12 (1-325) x US-09-591-435-11 (1-1212)

QY 69 LysAlaAlaValGlyGluLeuSerGluLysSerLysLeuGlnGluGly 120
: |||
DB 253 AAACCTGACTGCTGAGCTCTGAGCAAAATCAAGCTGTAGAGATCTATC 288

QY 89 ThrGlnLeuLysAlaAlaValGlyGluLeuProGluLysSerLys 103
: |||
DB 313 ACCCACTGAAGAGCTGTAGTAACTTACATCAAGAAATCTAAG 357

RESULT 6
: US-09-280-116-111
: Sequence 111, Application US/09280116A
: Patent No. 6331427
: GENERAL INFORMATION:
: APPLICANT: ROBINSON, Keith E.
: TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human
: FILE REFERENCE: 5800-24, 5800-2176965
: CURRENT APPLICATION NUMBER: US/09/280,116A
: CURRENT FILING DATE: 1999-04-26
: NUMBER OF SEQ ID NOS: 268
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 111
: LENGTH: 606
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: ADAM family of metalloproteinase
: NAME/KEY: misc_feature
: LOCATION: (1)..(606)
: OTHER INFORMATION: n = a, t, c or g
: US-09-280-116-111

Alignment Scores:
Pred. No.: 2.72          Length: 606
Score: 9.00             Matches: 9
Percent Similarity: 100.00%   Conservat: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 2.77%          Indels: 0
DB: 4                  Gaps: 0

US-09-831-458a-12 (1-325) x US-09-280-116-111 (1-606)

QY 34 AlaGlyValLeuValAlaAlaLeuVal 42
: |||
DB 199 GCTGAGCTGTGCTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCTGTGAG 225

RESULT 7
: US-08-836-443-2
: Sequence 2, Application US/08836443
: Patent No. 5884241
: GENERAL INFORMATION:
: APPLICANT: DOCHERTY, Andrew, J.P.
: APPLICANT: STOCOMBE, Patrick, M.
: TITLE OF INVENTION: DNA SEQUENCES CODING FOR A
: TITLE OF INVENTION: HUMAN METALLOPROTEINASE AND VARIAN
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, FORESTS & CUSMAN, LLC
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA

```


LENGTH: 3744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3744
PUBLICATION INFORMATION:
AUTHORS: Delisse-Gathoye, et al.
TITLE: Cloning, Partial Sequence, Expressions, and
TITLE: Antigenic Analysis of the Filamentous
TITLE: Hemagglutinin Gene of Bordetella Pertussis
JOURNAL: Infection and Immunity
VOLUME: 58
ISSUE: 9
PAGES: 2895-2905
DATE: September-1990
US-08-148-453-16

Alignment Scores:
Pred. No.: 152 Length: 3744
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 46% Indels: 0
DB: 2 Caps: 0

US-09-831-458a-12 (1-325) x US-08-348-353-16 (1-3744)
QY 46 ValLeuValAlaIaIeLeuValGln 43
DB 3291 GTACTGTAAGCATTCGTGTCAG 3268

RESULT 10
US-08-465-965-16/C
Sequence 16, Application US/08465965
Patent No. 5968512
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine
APPLICANT: Masure, Robert
TITLE OF INVENTION: Antibody Recognizing Endothelial Cell
TITLE OF INVENTION: Ligand for Leukocyte CR3
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,965
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/348,353
FILING DATE: 30-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,572
FILING DATE: 23-MAY-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/03725
FILING DATE: 04-MAY-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/695,613

FILING DATE: 03-MAY-1991
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, David
REGISTRATION NUMBER: 24,742
REFERENCE/DOCKET NUMBER: 600-1-097-1P111V2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 3744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3744
PUBLICATION INFORMATION:
AUTHORS: Delisse-Gathoye, et al.
TITLE: Cloning, Partial Sequence, Expressions, and
TITLE: Antigenic Analysis of the Filamentous
TITLE: Hemagglutinin Gene of Bordetella Pertussis
JOURNAL: Infection and Immunity
VOLUME: 58
ISSUE: 9
PAGES: 2895-2905
DATE: September-1990
US-08-465-965-16

Alignment Scores:
Pred. No.: 152 Length: 3744
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 46% Indels: 0
DB: 2 Caps: 0

US-09-831-458a-12 (1-325) x US-08-465-965-16 (1-3744)
QY 46 ValLeuValAlaIaIeLeuValGln 43
DB 3291 GTACTGTAAGCATTCGTGTCAG 3268

RESULT 11
US-08-465-966-16/C
Sequence 16, Application US/08465966
Patent No. 6015560
GENERAL INFORMATION:
APPLICANT: Tuomanen, Elaine
APPLICANT: Masure, Robert
TITLE OF INVENTION: Antibody recognizing Endothelial
TITLE OF INVENTION: Ligand for Leukocyte CR3
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,966
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/348,354
FILING DATE: 30-NOV-1994

```

CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/247,572
FILING DATE: 23-MAY-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/04725
FILING DATE: 04-MAY-1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/695,613
FILING DATE: 03-MAY-1991
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Jackson, David
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-097191D1V2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 3744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3744
PUBLICATION INFORMATION:
AUTHORS: Delisse-Gathoye, et al.
TITLE: Cloning, Partial Sequence, Expressions, and
TITLE: Antigenic Analysis of the Filamentous
TITLE: Hemagglutinin Gene of Bordetella Pertussis
JOURNAL: Infection and Immunity
VOLUME: 58
ISSUE: 9
PAGES: 2895-2905
DATE: September-1990
US-08-465-966-16

Alignment Scores:
Pred. No.: 152 Length: 3744
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DE: 3 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-465-966-16 (1-3744)
QY 36 ValdeValAlAlleuValgin 43
DB 3291 GTACTGTAGGAGATTCTGTGTCAG 3268

RESULT 12
US-08-474-077C-1
Sequence 1, Application US/08374077C
Patent No. 6027912
GENERAL INFORMATION:
APPLICANT: Hall, Linda M
APPLICANT: Ren, Dejian
APPLICANT: Zheng, Wei
APPLICANT: Dupald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha1
TITLE OF INVENTION: Calcium Channel Subunit
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWICKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA

```

```

TITLE: 22414-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,077C
FILING DATE: 19-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McCowan, Malcolm M.
REGISTRATION NUMBER: 39,700
REFERENCE/DOCKET NUMBER: 022650-264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703 836 6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 157..7704
US-08-374-077C-1

Alignment Scores:
Pred. No.: 313 Length: 8075
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: 3 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-374-077C-1 (1-8075)
QY 16 GAGTGTCTGGGAGGAGGATTA 23
DB 3153 GAGTGTCTGGGAGGAGGATTA 3176

RESULT 13
US-08-895-590-1
Sequence 1, Application US/08895590
Patent No. 6207410
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
APPLICANT: Ren, Dejian
APPLICANT: Zheng, Wei
APPLICANT: Dupald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWICKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
TITLE: 22414-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,590
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/374,888
FILING DATE: 19-JAN-1995

```

ATTORNEY/AGENT INFORMATION:
NAME: MCGOWAN, MALCOLM M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-263
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 157..7704
US-08-895-590-1

Alignment Scores:
Pred. No.: 313
Score: 8.00
Length: 8075
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatches: 0
Query Match: 2.468
Indels: 0
Gaps: 0
DB: 4

US-09-831-458A-12 (1-325) X US-08-895-590-1 (1-8075)

UY 16 C1GCTGCTTGGCGACGAGCATTA 23
|||||
DB 3153 GGGTGTCTTGGCGACGAGCATTA 3176

RESULT 14
US-09-539-879A-1
Sequence 1, Application US/09539879A
Patent No. 6436627
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
Ken, Dejian
Zheng, Wei
Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha-
Calcium Channel Subunit
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: HURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 609 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/539,879A
FILING DATE: 31-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/111,865
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/374,077
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MCGOWAN, MALCOLM M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 157..7704
US-09-539-879A-1

Alignment Scores:
Pred. No.: 313
Score: 8.00
Length: 8075
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatches: 0
Query Match: 2.468
Indels: 0
Gaps: 0
DB: 4

US-09-831-458A-12 (1-325) X US-09-539-879A-1 (1-8075)

UY 16 C1GCTGCTTGGCGACGAGCATTA 23
|||||
DB 3153 GGGTGTCTTGGCGACGAGCATTA 3176

RESULT 15
US-08-450-945-68
Sequence 68, Application US/08450945
Patent No. 5783383
GENERAL INFORMATION:
APPLICANT: Kondo, Kazuhito
Applicant: MocarSKI, Edward S. Jr.
TITLE OF INVENTION: LATENT TRANSCRIPTS AND PROMOTERS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dellinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,945
FILING DATE: 23-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Shultz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/DOCKET NUMBER: 8600-0157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: art isolate chr 3
FEATURE:
NAME/KEY: CDS

LOCATION: 1..132
US-08-450-945-68

Alignment Scores:
Pred. No.: 68 Length: 132
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458A-12 (1-325) X US-08-450-945-68 (1-132)
QY 35 G1YVAlIGnGInJenG1YLeu 41
|||||
DB 42 GGSGTTCCTGTCGAATCTTC 62

RESULT 16
US-08-976-161-68
Sequence 68, Application US/08976161
Patent No. 6194542

GENERAL INFORMATION:
APPLICANT: Kondo, Kazuhiro
APPLICANT: Mocariski, Edward S. Jr.
TITLE OF INVENTION: LATENT 1KANSCHIPS AND PROMOTERS
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: 350 Cambridge Avenue, Suite 250
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08-976,161
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/450,945
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sholtz, Charles K.
REGISTRATION NUMBER: 38,615
REFERENCE/WORK NUMBER: 8600-0157
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 132 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHEetical: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: antisense OPP 4
FEATURE:
NAME/KEY: CDS
LOCATION: 1..132
US-08-976-161-68

Alignment Scores:
Pred. No.: 68 Length: 132
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 4 Gaps: 0

US-09-831-458A-12 (1-325) X US-08-976-161-68 (1-132)
QY 35 G1YVAlIGnGInJenG1YLeu 41
|||||
DB 42 GGSGTTCCTGTCGAATCTTC 62

RESULT 17
US-07-743-518-23
Sequence 23, Application US/07743518
Patent No. 5397696

GENERAL INFORMATION:
APPLICANT: YANAGIHARA, RICHARD
APPLICANT: NEROKAR, VIVEK R.
APPLICANT: JENKINS, CAROL
APPLICANT: MILLER, MARK
APPLICANT: GARRETO, PALPH M.
TITLE OF INVENTION: PAPUA NEW GUINEA HUMAN T-LYMPHOTROPIC
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARV & CUSHMAN
STREET: 1615 L STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20046-5601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/743,518
FILING DATE: 19910812
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, MATSON T.
REGISTRATION NUMBER: 26,581
REFERENCE/WORK NUMBER: WTS/5683/84609/SAP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 28..443
US-07-743-518-23

Alignment Scores:
Pred. No.: 178 Length: 369
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458A-12 (1-325) X US-07-743-518-23 (1-369)
QY 8 ArvVAlIGnGInJenG1YLeu 14
|||||
DB 60 AGAGTACAGCAGGCTG 40

RESULT 18
US-09-280-116-105
Sequence 105, Application US/09280116A
Patent No. 6331427
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
FILE REFERENCE: 5800-24, 035800/176965
CURRENT APPLICATION NUMBER: US/09/280,116A
CURRENT FILING DATE: 1999-03-26
NUMBER OF SEQ ID NOS: 268
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 105
LENGTH: 407
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: trypsin-like serine proteases
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)-(407)
OTHER INFORMATION: n = a, t, c or g
US-09-280-116-105

Alignment Scores:
Pred. No.: 195 Length: 407
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 4 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-280-116-105 (1-407)

OY 10 GlnGlnLeuGlyLeuGly 16
|||||
DB 114 CAGCACTGGGCTTACTAGGG 134

RESULT 19
US-07-662-198B-1/C
Sequence 1, Application US/07662198B
Patent No. 5262528
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
APPLICANT: Lee, Sam W.
APPLICANT: Tomasetto, Catherine
TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/662,198B
FILING DATE: 19910228
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/049001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 452
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
US-07-662-198B-1

Alignment Scores:
Pred. No.: 215 Length: 452
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458a-12 (1-325) x US-07-662-198B-1 (1-452)

OY 32 LeuLeuAlaGlyValLeuVal 48
|||||
DB 116 CTCTGGCAGCAGTCTGTG 96

RESULT 20
US-08-322-742-1/C
Sequence 1, Application US/08322742
Patent No. 5688641
GENERAL INFORMATION:
APPLICANT: Sager, Ruth
TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 502 or 555X
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/322,742
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/948,848
FILING DATE: September 1, 1992
APPLICATION NUMBER: 07/844,296
FILING DATE: February 28, 1992
APPLICATION NUMBER: 07/552,216
FILING DATE: February 28, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/048003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 452
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-322-742-1
Alignment Scores:
Pred. No.: 215 Length: 452
Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.15% Indels: 0
 DB: 1 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-322-742-1 (1-452)

UY 32 LepteqalaglyValleuval 38
 DB 116 CTTCTCCGACGACTCTTCG 96

RESULT 21

US-08-450-945-66
 : Sequence 66, Application US/08450945
 : Patent No. 5783383

: GENERAL INFORMATION:

: APPLICANT: Kondo, Kazuhiko

: APPLICANT: Mocarski, Edward S. Jr.

: TITLE OF INVENTION: OF CYTOMEGALOVIRUS

: NUMBER OF SEQUENCES: 75

: CORRESPONDENCE ADDRESS:

: ADDRESS: Deninger & Associates

: STREET: 350 Cambridge Avenue, Suite 250

: CITY: Palo Alto

: STATE: CA

: COUNTRY: USA

: ZIP: 94306

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent In Release #1 0, Version #1 30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/450,945

: FILING DATE: 23-MAY-1995

: CLASSIFICATION: 424

: ATTORNEY/AGENT INFORMATION:

: NAME: Shultz, Charles K

: REGISTRATION NUMBER: 38,615

: REFERENCE/DOCKET NUMBER: 8600-0157

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (415) 324-0880

: TELEFAX: (415) 324-0960

: INFORMATION FOR SEQ ID NO: 66:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 462 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: double

: TOPOLOGY: linear

: MOLECULE TYPE: cDNA to mRNA

: HYPOTHETICAL: NO

: ANTI-SENSE: NO

: ORIGINAL SOURCE:

: INDIVIDUAL ISOLATE: antisense ORF 2

: FEATURE:

: NAME/KEY: CDS

: LOCATION: 1..462

: US-08-450-945-66

Alignment Scores:
 Pred. No.: 219 Length: 462
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.15% Indels: 0
 DB: 1 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-450-945-66 (1-462)

UY 35 CiyValleuValAIdJleu 41
 DB 67 GGGTTTCGTCGCAATCTCG 87

RESULT 22

US-08-976-161-66

: Sequence 66, Application US/08976161

: Patent No. 6194542

: GENERAL INFORMATION:

: APPLICANT: Kondo, Kazuhiko

: APPLICANT: Mocarski, Edward S. Jr.

: TITLE OF INVENTION: LATENT TRANSCRIPTS AND PROMOTERS

: NUMBER OF SEQUENCES: 75

: CORRESPONDENCE ADDRESS:

: ADDRESS: Deninger & Associates

: STREET: 350 Cambridge Avenue, Suite 250

: CITY: Palo Alto

: STATE: CA

: COUNTRY: USA

: ZIP: 94306

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent In Release #1 0, Version #1 30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/976,161

: FILING DATE:

: CLASSIFICATION:

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: 08/450,945

: FILING DATE:

: ATTORNEY/AGENT INFORMATION:

: NAME: Shultz, Charles K

: REGISTRATION NUMBER: 38,615

: REFERENCE/DOCKET NUMBER: 8600-0157

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (415) 324-0880

: TELEFAX: (415) 324-0960

: INFORMATION FOR SEQ ID NO: 66:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 462 base pairs

: TYPE: nucleic acid

: STRANDEDNESS: double

: TOPOLOGY: linear

: MOLECULE TYPE: cDNA to mRNA

: HYPOTHETICAL: NO

: ANTI-SENSE: NO

: ORIGINAL SOURCE:

: INDIVIDUAL ISOLATE: antisense ORF 2

: FEATURE:

: NAME/KEY: CDS

: LOCATION: 1..462

: US-08-976-161-66

Alignment Scores:
 Pred. No.: 219 Length: 462
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.15% Indels: 0
 DB: 4 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-976-161-66 (1-462)

UY 35 CiyValleuValAIdJleu 41
 DB 67 GGGTTTCGTCGCAATCTCG 87

RESULT 23

US-08-458-084-5

: Sequence 5, Application US/08458084

: Patent No. 5624837

: GENERAL INFORMATION:

: APPLICANT: Podor, William L

```

: APPLICANT: Rollins, Scott
: APPLICANT: Squinto, Stephen P
: TITLE OF INVENTION: Chimeric Complement
: TITLE OF INVENTION: Inhibitor Proteins
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Maurice M. Klee
: STREET: 1951 Burr Street
: CITY: Fairfield
: STATE: Connecticut
: COUNTRY: USA
: ZIP: 06430
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 750 Kb storage
: COMPUTER: Dell 486/50
: OPERATING SYSTEM: DOS 6.2
: SOFTWARE: Wordperfect 6.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/458,084
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Klee, Maurice M.
: REGISTRATION NUMBER: 30,399
: REFERENCE/DOCKET NUMBER: ALX-120
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (203) 254-1400
: TELEFAX: (203) 254-1101
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 469 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Double
: TOPOLOGY: Linear
: MOLECULE TYPE: cDNA to mRNA
: DESCRIPTION: AGCIP full length cDNA
: HYPOTHEITICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Cercopithecus aethiops
: CELL LINE: COS-1 (ATCC CRL 1650)
: US-08-458-084-5

Alignment Scores:
Pred. No.: 223 Length: 469
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
Gaps: 0

US-09-831-458A-12 (1-325) x US-08-458-084-5 (1-469)

Oy 47 ValProserSerLeuSerGln 53
Db 383 GTGCTTATCCCTAAGTCAA 403

RESULT 24
US-08-205-508-5
: Sequence 5, Application US/08205508
: Patent No. 5627264
: GENERAL INFORMATION:
: APPLICANT: Fodor, William L
: APPLICANT: Rollins, Scott
: APPLICANT: Squinto, Stephen P
: TITLE OF INVENTION: Chimeric Complement
: TITLE OF INVENTION: Inhibitor Proteins
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Maurice M. Klee
: STREET: 1951 Burr Street
: CITY: Fairfield
: STATE: Connecticut

```

```

: COUNTRY: USA
: ZIP: 06430
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 750 Kb storage
: COMPUTER: Dell 486/50
: OPERATING SYSTEM: DOS 6.2
: SOFTWARE: Wordperfect 6.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/205,508
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Klee, Maurice M.
: REGISTRATION NUMBER: 30,399
: REFERENCE/DOCKET NUMBER: ALX-120
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (203) 254-1400
: TELEFAX: (203) 254-1101
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 469 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Double
: TOPOLOGY: Linear
: MOLECULE TYPE: cDNA to mRNA
: DESCRIPTION: AGCIP full length cDNA
: HYPOTHEITICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Cercopithecus aethiops
: CELL LINE: COS-1 (ATCC CRL 1650)
: US-08-205-508-5

Alignment Scores:
Pred. No.: 223 Length: 469
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
Gaps: 0

US-09-831-458A-12 (1-325) x US-08-205-508-5 (1-469)

Oy 47 ValProserSerLeuSerGln 54
Db 383 GTGCTTATCCCTAAGTCAA 403

RESULT 25
US-08-482-148-2
: Sequence 2, Application US/082148
: Patent No. 5847082
: GENERAL INFORMATION:
: APPLICANT: Rother, Russell
: APPLICANT: Rollins, Scott
: APPLICANT: Squinto, Stephen P
: TITLE OF INVENTION: Terminal Complement
: TITLE OF INVENTION: Inhibitor Fusion Genes and Proteins
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Seth A. Fidler
: STREET: Alexion Pharmaceuticals, 45 Science Park
: CITY: New Haven
: STATE: Connecticut
: COUNTRY: USA
: ZIP: 06511
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.4 Mb storage
: COMPUTER: IBM compatible (Pentium)
: OPERATING SYSTEM: Windows 95 under MS D.S
: SOFTWARE: Microsoft Word for Windows
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/482,148
: FILING DATE: 6/07/95

```

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fidel, Seth A.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-129.1 Div
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776 1790
TELEFAX: (203) 772 3655
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: AGMCIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cercopithecus aethiops
CELL LINE: COS-1 (ATCC CRL 1650)
US-08-482-148-2

Alignment Scores:
Pred. No.: 223 Length: 469
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: Gaps: 0

US-09-831-458a-12 (1-325) x US-08-482-148-2 (1-469)

Cy 47 ValProSerSerLeuSerGln 53
|||||
Db 383 GTGCTTCATCCTAGTCA 403

RESULT 26
FCT-US95-02944-2
Sequence 2, Application PC/TUS9502944
GENERAL INFORMATION:
APPLICANT: Rother, Russell
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Terminal Complement
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: Word Perfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02944
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/295,720
FILING DATE: 3-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-129PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400

TELEFAX: (203) 254 1101
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: AGMCIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cercopithecus aethiops
CELL LINE: COS-1 (ATCC CRL 1650)
FCT US95 02944 2

Alignment Scores:
Pred. No.: 223 Length: 469
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: Gaps: 0

US-09-831-458a-12 (1-325) x PCT-US95-02944-2 (1-469)

Cy 47 ValProSerSerLeuSerGln 53
|||||
Db 383 GTGCTTCATCCTAGTCA 403

RESULT 27
FCT-US95-02945-5
Sequence 5, Application PC/TUS9502945
GENERAL INFORMATION:
APPLICANT: Podor, William L.
APPLICANT: Rother, Russell
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Chimeric Complement
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 720 kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/205,508
FILING DATE: 3-MAR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-120PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255-1400
TELEFAX: (203) 254-1101
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear

MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: AGC1P full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Cercopithecus aethiops
CELL LINE: COS-1 (ATCC CRL 1650)
PCT-US95-02945-5

Alignment Scores:

Pred. No.:	223	Length:	469
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.15%	Indels:	0
DB:	5	Gaps:	0

US-09-831-458A-12 (1-325) x PCT-US95-02945-5 (1-469)

UY 47 ValProserleusergin 53

Db 483 GGCTTTCATCTAACTGAA 403

RESULT 28

US-08-031-148-1

Sequence 1, Application US/08031148

Patent No. 5424398

GENERAL INFORMATION:

APPLICANT: Middeldorp, Jaap Michiel.

TITLE OF INVENTION: Peptides and nucleic acid sequences

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Akzo Pharma

STREET: 1330-A Piccard Drive

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850-4377

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/031,148

FILING DATE: 19930312

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 92200721.6

FILING DATE: 13-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Bobrowicz, Donna

REGISTRATION NUMBER: 32,196

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 538 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Epstein-Barr virus

US-08-031-148-1

Alignment Scores:

Pred. No.:	253	Length:	538
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.15%	Indels:	0
DB:	1	Gaps:	0

US-09-831-458A-12 (1-325) x US-08-041-148-1 (1-538)

QY 13 GYLCUCLUGCYCYSLUGCY 14

Db 192 GGACTTTTGCGTCTCGGC 212

RESULT 29

US-08-415-838-1

Sequence 1, Application US/08415848

Patent No. 6008427

GENERAL INFORMATION:

APPLICANT: Middeldorp, Jaap Michiel.

TITLE OF INVENTION: Peptides and nucleic acid sequences

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Akzo-No. 6008427e1 Patent Department

STREET: 1300 Piccard Drive, Suite 206

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/415,848

FILING DATE: 03-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 92200721.6

FILING DATE: 13-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Gormley, Mary E.

REGISTRATION NUMBER: 34,404

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 538 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Epstein-Barr virus

US-08-415-838-1

Alignment Scores:

Pred. No.:	253	Length:	538
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.15%	Indels:	0
DB:	3	Gaps:	0

US-09-831-458A-12 (1-325) x US-08-415-838-1 (1-538)

QY 13 GYLCUCLUGCYCYSLUGCY 14

Db 192 GGACTTTTGCGTCTCGGC 212

RESULT 30

US-09-205-169-1

Sequence 1, Application US/09205169

Patent No. 6365717

GENERAL INFORMATION:

APPLICANT: Middeldorp, Jaap Michiel.

TITLE OF INVENTION: Peptides and nucleic acid sequences

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Akzo-No. 6465717e1 Patent Department

STREET: 1300 Piccard Drive, Suite 206


```

: NAME: Agriis Dr., Cheryl H.
: REGISTRATION NUMBER: 34,086
: REFERENCE/DOCKET NUMBER: 3764, 400-US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-867-0123
: TELEFAX: 212-878-9655
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 679 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-459-871-9

Alignment Scores:
Pred. No.: 315 Length: 679
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 1 Gaps: 0

US-09-831-458A-12 (1-325) x US-08-459-871-9 (1-679)

OY 110 GlutathioneS-transferase 116
DB 296 GACCTGACGATTCGAAAGCA 276

RESULT 36
US-09-419-568F-24/c
: Sequence 24, Application US/09419568F
: Patent No. 6331613
: GENERAL INFORMATION:
: APPLICANT: Dumoutier, Laure
: APPLICANT: Louhed, Jamila
: TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fas
: FILE REFERENCE: LUD 5543.2
: CURRENT APPLICATION NUMBER: US/09/419,568F
: PRIOR FILING DATE: 1999-10-18
: PRIOR APPLICATION NUMBER: US09/354,243
: PRIOR FILING DATE: 1999-07-16
: PRIOR APPLICATION NUMBER: US09/178,973
: PRIOR FILING DATE: 1998-10-26
: NUMBER OF SEQ ID NOS: 29
: SEQ ID NO 24
: LENGTH: 690
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: US-09-419-568F-24

Alignment Scores:
Pred. No.: 319 Length: 690
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 4 Gaps: 0

US-09-831-458A-12 (1-325) x US-09-419-568F-24 (1-690)

OY 271 ScFvProteinScFvProtein 277
DB 550 TCTCCACTCTCTCCACGCTT 530

RESULT 37
US-09-354-243B-24/c
: Sequence 24, Application US/09354243B
: Patent No. 6359117
: GENERAL INFORMATION:
: APPLICANT: Dumoutier, Laure
```

```

: APPLICANT: Louhed, Jamila
: APPLICANT: Renaud, Jean-Christophe
: TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible
: FILE REFERENCE: (TIFS) The Proteins Encoded, and Uses Thereof
: CURRENT APPLICATION NUMBER: US/09/354,243
: PRIOR FILING DATE: 1999-07-16
: PRIOR APPLICATION NUMBER: US09/178,973
: PRIOR FILING DATE: 1998-10-26
: NUMBER OF SEQ ID NOS: 29
: SEQ ID NO 24
: LENGTH: 690
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: US-09-354-243B-24

Alignment Scores:
Pred. No.: 319 Length: 690
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 4 Gaps: 0

US-09-831-458A-12 (1-325) x US-09-354-243B-24 (1-690)

OY 271 ScFvProteinScFvProtein 277
DB 550 TCTCCACTCTCTCCACGCTT 530

RESULT 38
US-08-458-084-4
: Sequence 4, Application US/08458084
: Patent No. 5624847
: GENERAL INFORMATION:
: APPLICANT: Fodor, William L.
: APPLICANT: Hollins, Scott
: APPLICANT: Squitto, Stephen P.
: TITLE OF INVENTION: Chimeric Complement
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Maurice M. Klee
: STREET: 1951 Burr Street
: CITY: Fairfield
: STATE: Connecticut
: COUNTRY: USA
: ZIP: 06430
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 750 Kb storage
: COMPUTER: Dell 486/50
: OPERATING SYSTEM: DOS 6.2
: SOFTWARE: WordPerfect 6.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/458,084
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Klee, Maurice M.
: REGISTRATION NUMBER: 30,399
: REFERENCE/DOCKET NUMBER: ALX-120
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (203) 255-1400
: TELEFAX: (203) 254-1101
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 763 base pairs
: TYPE: Nucleic Acid
: STRANDEDNESS: Double
: TOPOLOGY: Linear
: MOLECULE TYPE: cDNA to mRNA
```

DESCRIPTION: BACIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Papio hamadryas
IMMEDIATE SOURCE:
LIBRARY: Baboon Spleen Lambda ZAP11 cDNA
LIBRARY: Library, Catalog # 936103,
LIBRARY: Stratagene Cloning Systems,
LIBRARY: La Jolla, California
US-08-458-084-4

Alignment Scores:
Pred. No.: 351
Score: 7.00
Percent Similarity: 100.00%
Best local Similarity: 100.00%
Query Match: 2.15%
DB: 1
Length: 763
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x US-08-458-084-4 (1-763)

OY 47 ValProSerLeuSerGln 53
|||||
DB 506 GTGCTTCATCCTAAGTCAA 526

RESULT 39
US-08-205-508-4
Sequence 4, Application US/08205508
Patent No. 5627264
GENERAL INFORMATION:
APPLICANT: Fodor, William L
APPLICANT: Rollins, Scott
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Chimeric Complement
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 750 Kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/205,508
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-120
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255-1400
TELEFAX: (203) 254-1101
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: BACIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Papio hamadryas
IMMEDIATE SOURCE:
LIBRARY: Baboon Spleen Lambda ZAP11 cDNA
LIBRARY: Library, Catalog # 936103,
LIBRARY: Stratagene Cloning Systems, La Jolla

LIBRARY: Baboon Spleen Lambda ZAP11 cDNA
LIBRARY: Library, Catalog # 936103,
LIBRARY: Stratagene Cloning Systems,
LIBRARY: La Jolla, California
US-08-205-508-4

Alignment Scores:
Pred. No.: 351
Score: 7.00
Percent Similarity: 100.00%
Best local Similarity: 100.00%
Query Match: 2.15%
DB: 1
Length: 763
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x US-08-205-508-4 (1-763)

OY 47 ValProSerLeuSerGln 53
|||||
DB 506 GTGCTTCATCCTAAGTCAA 526

RESULT 40
US-08-482-148-1
Sequence 1, Application US/08482148
Patent No. 5847082
GENERAL INFORMATION:
APPLICANT: Rother, Russell
APPLICANT: Rollins, Scott
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Terminal Complement
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: Alexion Pharmaceuticals, 25 Science Park
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4 Mb storage
COMPUTER: IBM compatible (pentium)
OPERATING SYSTEM: Windows 95 under MS DOS
SOFTWARE: Microsoft Word for Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,148
FILING DATE: 6/07/95
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fidel, Seth A.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-120,1 Div
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-3655
TELEFAX: (203) 776-1790
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Linear
MOLECULE TYPE: cDNA to mRNA
DESCRIPTION: BACIP full length cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Papio hamadryas
IMMEDIATE SOURCE:
LIBRARY: Baboon Spleen Lambda ZAP11 cDNA
LIBRARY: Library, Catalog # 936103,
LIBRARY: Stratagene Cloning Systems, La Jolla

US-08-482-148-1

Alignment Scores:
Pred. No.: 351
Length: 763

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 2 Gaps: 0

US-09-831-458a-12 (1-325) x US-08-482-148-1 (1-763)

QY 47 ValProSeSeSeSeSeSeSeSeSe 53

DB 506 GTGCTTCATCCTCAAGTCA 526

RESULT 41

PCT-US95-02944-1
Sequence 1, Application PC/TUS9502944

GENERAL INFORMATION:

APPLICANT: Kother, Russell
APPLICANT: Rollins, Scott
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Terminal Complement
TITLE OF INVENTION: Inhibitor Fusion Genes and Proteins
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA

ZIP: 06430

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 720 Kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: Word Perfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02944

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/7205, 720

FILING DATE: 3-MAR-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Klee, Maurice M.

REGISTRATION NUMBER: 30,399

REFERENCE/DOCKET NUMBER: ALX-129PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (203) 255-1400

TELEFAX: (203) 254-1101

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 763 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Double

TOPOLOGY: Linear

MOLECULE TYPE: cDNA to mRNA

DESCRIPTION: BABCP full length cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Papio hamadryas

IMMEDIATE SOURCE:

LIBRARY: Baboon Spleen Lambda ZapII cDNA Library, Catalog

LIBRARY: #936103, Stratagene Cloning Systems, La Jolla, California

PCT-US95-02944-1

Alignment Scores:

Pred. No.: 351 Length: 763
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 5 Gaps: 0

US-09-831-458a-12 (1-325) x PCT US95-02944-1 (1-763)

QY 47 ValProSeSeSeSeSeSeSeSeSe 53

DB 506 GTGCTTCATCCTCAAGTCA 526

RESULT 42

PCT-US95-02945-4

Sequence 4, Application PC/TUS9502945

GENERAL INFORMATION:

APPLICANT: Fodor, William L.
APPLICANT: Rollins, Scott
APPLICANT: Squinto, Stephen P
TITLE OF INVENTION: Chimeric Complement
TITLE OF INVENTION: Inhibitor Proteins
NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA

ZIP: 06430

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 750 Kb storage
COMPUTER: Dell 486/50
OPERATING SYSTEM: DOS 6.2
SOFTWARE: WordPerfect 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02945

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/205,508

FILING DATE: 3-MAR-1994

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Klee, Maurice M.

REGISTRATION NUMBER: 30,399

REFERENCE/DOCKET NUMBER: ALX-120PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (203) 255-1400

TELEFAX: (203) 254-1101

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 763 base pairs

TYPE: Nucleic Acid

STRANDEDNESS: Double

TOPOLOGY: Linear

MOLECULE TYPE: cDNA to mRNA

DESCRIPTION: BABCP full length cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Papio hamadryas

IMMEDIATE SOURCE:

LIBRARY: Baboon Spleen Lambda ZapII cDNA Library, Catalog

LIBRARY: #936103, Stratagene Cloning Systems, La Jolla, California

PCT-US95-02945-4

Alignment Scores:

Pred. No.: 351 Length: 763
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 5 Gaps: 0

US-09-831-458a-12 (1-325) x PCT US95-02945-4 (1-763)

QY 47 ValProSeSeSeSeSeSeSeSeSe 53

DB 506 GTGCTTCATCCTCAAGTCA 526

Db 506 GTCCTGAGCTCTTCCCTTC 99

RESULT 43
US-09-051-696-5/c
Sequence 5, Application US/09051696
Patent No. 6319495
GENERAL INFORMATION:
APPLICANT: Pollock, Allan S.
Christakos, Sylvia
Reddy, Daphne
TITLE OF INVENTION: Method for Restoring Glucose Responsiveness to Insulin Secretion
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Fifth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/051.696
FILING DATE: 17-Apr-1998
CLASSIFICATION: <unknown>
APPLICATION NUMBER: US 60/005.386
FILING DATE: 10-OCT-1995
APPLICATION NUMBER: WO PCT/US96/16746
FILING DATE: 17-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Kenneth A.
REGISTRATION NUMBER: 31.677
REFERENCE/DOCKET NUMBER: 02307E-062710US
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 791 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-051-696-5

Alignment Scores:
Pred. No.: 363 Length: 791
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x US-09-051-696-5 (1-791)

Db 119 GTCCTGAGCTCTTCCCTTC 99

RESULT 44
US-09-416-509C-2/c
Sequence 2, Application US/09416509C
Patent No. 6469141
GENERAL INFORMATION:
APPLICANT: Bergson, Claire
TITLE OF INVENTION: D1-like Dopamine Receptor Activity Modifying Protein
FILE REFERENCE: M0351-205012

US-09-831-458a-12 (1-325) x US-09-416-509C-2 (1-936)

Db 115 GTCCTGAGCTCTTCCCTTC 99

RESULT 45
US-08-973-068-28
Sequence 28, Application US/08973068
Patent No. 6127604
GENERAL INFORMATION:
APPLICANT: Dale, James Langham
APPLICANT: Harding, Robert Maxwell
APPLICANT: Dugdale, Benjamin
APPLICANT: Beetham, Peter Ronald
APPLICANT: Haefl, Gregory John
TITLE OF INVENTION: INTERGENIC REGIONS OF BANANA BUNCHY DISEASE
FILE REFERENCE: 09657/002001
CURRENT APPLICATION NUMBER: US/08/973.068
CURRENT FILING DATE: 1998-03-12
EARLIER APPLICATION NUMBER: PCT/AU96/00345
EARLIER FILING DATE: 1996-05-31
NUMBER OF SEQ ID NOS: 61
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 982
TYPE: DNA
ORGANISM: Banana Bunchy Top Virus (BBTV)
US-08-973-068-28

Alignment Scores:
Pred. No.: 444 Length: 982
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
Gaps: 0

US-09-831-458a-12 (1-325) x US-08-973-068-28 (1-982)

Db 941 TCGAAGTCCTTCGAGTTT 961

Db 941 TCGAAGTCCTTCGAGTTT 961

Search completed: December 7, 2002, 14:54:28
Job time : 81 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Computer 113

UM protein - nucleic search, using frame_plus_f2n model

Run on: December 7, 2002, 12:27:12, Search time: 61 seconds

(without alignments)
2077.510 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325
Sequence: 1 MSDSKFRVQGLIAPGLGH PCTVNWYVCKKPAKFPNP 325

Scoring table: OLIGO

Xgapop 60.0, Ygapext 60.0
Ygapop 60.0, Ygapext 60.0
Delop 6.0, Delext 7.0

Searched: 350425 seqs, 194966369 residues

Word size: 5

Total number of hits satisfying chosen parameters: 43445

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_model -DEV=xib
-O=/cgn2.1/USP70/US09831458/rmat.05122002.1041416803/rff_jquery.fasta_1_13
-DB=Published.Applications_USA_QFMT-fastlap -SUFFIX=ol1.rmpb -MINMATCH=0 1
-LOOKUP=0 -LOOKEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS-human40.cdi -LIST=45 -DUALIGN=200 -THR_SCORE=quality -THR_MIN=5
-ALIGN=45 -MODE=LOCAL -OUTEXT=ptc -NORM=ext -HEAPS17E=500 -MINFN=0
-MAXLEN=200000000 -USER=US09831458 -ACGN.1.1.21 -runat.05122002.1041416803
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGOUDRY -NRC_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: Published.Applications_USA:

1: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
3: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
4: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
5: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
6: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
7: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
8: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
9: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
10: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
11: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
12: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
13: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:
14: /cgn2.6/ptodata/2/pubpna/US07_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	Query Match length	DB ID	Description
1	51	15 7	592 10	US-09-864-761-7350 Sequence 7350, Ap
2	50	15 4	152 10	US-09-864-761-24072 Sequence 24072, A
3	9	2 8	1183 10	US-09-983-531A-15 Sequence 15, Appl
4	9	2 8	2404 10	US-09-983-531A-25 Sequence 25, Appl

5	9	2 8	2669 10	US-09-983-531A-19	Appl
6	9	2 8	2757 9	US-10-020-743-2	Appl
7	9	2 8	2757 9	US-10-155-470-2	Appl
8	9	2 8	2757 9	US-10-155-470-2	Appl
9	9	2 8	2781 9	US-10-020-743-1	Appl
10	9	2 8	2868 9	US-10-020-743-7	Appl
11	9	2 8	2892 9	US-10-020-743-5	Appl
12	9	2 8	3512 9	US-10-020-743-9	Appl
13	8	2 5	623 10	US-09-894-638A-110	Appl
14	8	2 5	884 10	US-09-770-445-515	Appl
15	8	2 5	893 10	US-09-894-638A-95	Appl
16	8	2 5	952 10	US-09-770-445-328	Appl
17	8	2 5	1146 10	US-09-887-576-673	Appl
18	8	2 5	2006 10	US-09-887-576-673	Appl
19	8	2 5	3150 10	US-09-962-436-131	Appl
20	8	2 5	3150 10	US-09-962-436-131	Appl
21	8	2 5	3150 10	US-09-962-436-131	Appl
22	8	2 5	3150 10	US-09-962-436-131	Appl
23	8	2 5	18871 10	US-09-962-436-131	Appl
24	7	2 2	124 10	US-09-962-436-131	Appl
25	7	2 2	128 10	US-09-962-436-131	Appl
26	7	2 2	128 10	US-09-962-436-131	Appl
27	7	2 2	128 10	US-09-962-436-131	Appl
28	7	2 2	128 10	US-09-962-436-131	Appl
29	7	2 2	128 10	US-09-962-436-131	Appl
30	7	2 2	128 10	US-09-962-436-131	Appl
31	7	2 2	128 10	US-09-962-436-131	Appl
32	7	2 2	128 10	US-09-962-436-131	Appl
33	7	2 2	128 10	US-09-962-436-131	Appl
34	7	2 2	128 10	US-09-962-436-131	Appl
35	7	2 2	128 10	US-09-962-436-131	Appl
36	7	2 2	128 10	US-09-962-436-131	Appl
37	7	2 2	128 10	US-09-962-436-131	Appl
38	7	2 2	128 10	US-09-962-436-131	Appl
39	7	2 2	128 10	US-09-962-436-131	Appl
40	7	2 2	128 10	US-09-962-436-131	Appl
41	7	2 2	128 10	US-09-962-436-131	Appl
42	7	2 2	128 10	US-09-962-436-131	Appl
43	7	2 2	128 10	US-09-962-436-131	Appl
44	7	2 2	128 10	US-09-962-436-131	Appl
45	7	2 2	128 10	US-09-962-436-131	Appl

ALIGNMENTS

RESULT 1
US-09-864-761-7350
Sequence 7350, Application US/09864761
Patent No US/200408763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLE
FILE REFERENCE: Accolma-X-1
CURRENT APPLICATION NUMBER: US/09864761
PRIORITY FILING DATE: 2001-05-23
PRIORITY FILING DATE: 2000-02-04
PRIORITY FILING DATE: 2000-02-04
PRIORITY FILING DATE: 2000-05-26
PRIORITY FILING DATE: 2000-08-03
PRIORITY FILING DATE: 2000-10-04
PRIORITY FILING DATE: 2000-09-27
PRIORITY FILING DATE: 2000-09-27
PRIORITY FILING DATE: 2001-01-30
PRIORITY FILING DATE: 2001-01-30

[illegible]

```
Db      63   ICTAACCCAGGGAACCTGTGCATCTCGTCAACGGCTGGCAGAAGTCAGGCCGGCAG    123
QY      229   LeuValIleLeuThrAlaGluGln     238
          |||||
Db      123   CTCTCTTAATCAAACATGCTGTAAGAGTAG    152

RESULT 3
US-09-983-531A-15
: Sequence 15, Application US/09983531A
: Patent No. US20020147132A1
: GENERAL INFORMATION:
: APPLICANT: Fujisawa, Atsuko
: APPLICANT: Yamakawa, Toru
: APPLICANT: Shirakawa, Kamon
: APPLICANT: Chitose, Orif
: APPLICANT: Ogawa, Naoki
: TITLE OF INVENTION: Meltrins
: FILE REFERENCE: 11-22-99 sequence submission
: CURRENT APPLICATION NUMBER: US/09/983,531A
: PRIOR FILING DATE: 2001-10-24
: PRIOR APPLICATION NUMBER: JP 8-61756
: PRIOR FILING DATE: 1996-02-23
: PRIOR APPLICATION NUMBER: PCT/JP96/03017
: PRIOR FILING DATE: 1996-10-17
: NUMBER OF SEQ ID NOS: 44
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 15
: LENGTH: 1183
: TYPE: DNA
: ORGANISM: Unknown Organism
: FEATURE:
? OTHER INFORMATION: Description of Unknown Organism: Human meltrin
? OTHER INFORMATION: beta derived from cDNA
? NAME/KEY: CDS
? LOCATION: (2)..(1183)
US-09-983-531A-15

Alignment Scores:
Prod No.:           3.22             Length:         1183
Score:              9.00            Matches:        9
Percent Similarity: 100.00%         Conservative:   9
Best local Similarity: 100.00%       Mismatches:    0
Query Match:        2.77%            Indels:        0
DB:                 10               Gaps:         0

US-09-831-458A-12 (1-325) x US-09-983-531A-15 (1-1183)
QY      34   AlaglyValIleValAlaIleVal     42
          ||||||
Db      998   GTTGAGTTGTTGGCATCTTGATTG    1024

RESULT 4
US-09-983-531A-25
: Sequence 25, Application US/09983531A
: Patent No. US20020147132A1
: GENERAL INFORMATION:
: APPLICANT: Fujisawa, Atsuko
: APPLICANT: Yamakawa, Toru
: APPLICANT: Shirakawa, Kamon
: APPLICANT: Chitose, Orif
: APPLICANT: Ogawa, Naoki
: TITLE OF INVENTION: Meltrins
: FILE REFERENCE: 11-22-99 sequence submission
: CURRENT APPLICATION NUMBER: US/09/983,531A
: CURRENT FILING DATE: 2001-10-24
: PRIOR APPLICATION NUMBER: JP 8-61756
: PRIOR FILING DATE: 1996-02-23
: PRIOR APPLICATION NUMBER: PCT/JP96/03017
: PRIOR FILING DATE: 1996-10-17
: NUMBER OF SEQ ID NOS: 44
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 25
```

```

1  LENGTH: 2404
2  TYPE: DNA
3  ORGANISM: Unknown Organism
4  FEATURE:
5  OTHER INFORMATION: Description of Unknown Organism: (1)
6  OTHER INFORMATION: JM109 (pM1 beta 247), human meltrin b
7  NAME/KEY: CDS
8  LOCATION: (1) .. (1017)
9  US-09-983-531A-25
10
11 Alignment Scores:
12 Pred. No.: 6.38
13 Score: 9.00
14 Percent Similarity: 100.00%
15 Best local Similarity: 100.00%
16 Query Match: 2.77%
17 DB: 10
18
19 US-09-831-458A-12 (1-325) x US-09-983-531A-25 (1-2404)
20
21 QY 34 AlaGlyValIleuValAlaIleuVal 42
22 ||||||||||||||||||||||||
23 DB 735 GCTGAGTGTGTTGTCGCACTTGGTG 759
24
25 RESULT 5
26 US-09-983-531A-19
27 Sequence 19, Application US/09984541A
28 Patent No. US20020147132A1
29 GENERAL INFORMATION:
30 APPLICANT: Fujisawa, Atsuko
31 APPLICANT: Yamakawa, Toru
32 APPLICANT: Shirakawa, Kanon
33 APPLICANT: Chitose, Oriti
34 APPLICANT: Ogawa, Naoki
35 TITLE OF INVENTION: Meltrins
36 FILE REFERENCE: 11-22-99 sequence submission
37 CURRENT APPLICATION NUMBER: 02/097483, 531A
38 CURRENT FILING DATE: 2001-10-24
39 PRIOR APPLICATION NUMBER: JP 0-61756
40 PRIOR FILING DATE: 1995-02-24
41 PRIOR APPLICATION NUMBER: 11-11703017
42 PRIOR FILING DATE: 1996-10-17
43 NUMBER OF SEQ. ID NOS: 44
44 SOFTWARE: PatentIn Ver. 2.1
45 SEQ. ID NO 19
46 LENGTH: 2669
47 TYPE: DNA
48 ORGANISM: Unknown Organism
49 FEATURE:
50 OTHER INFORMATION: Description of Unknown Organism: (1)
51 OTHER INFORMATION: JM109 (pM1 beta 247) and JM109 (pM1 beta
52 NAME/KEY: CDS
53 LOCATION: (2) .. (1282)
54 US-09-983-531A-19
55
56 Alignment Scores:
57 Pred. No.: 7.06
58 Score: 9.00
59 Percent Similarity: 100.00%
60 Best local Similarity: 100.00%
61 Query Match: 2.77%
62 DB: 10
63
64 US-09-831-458A-12 (1-325) x US-09-983-531A-19 (1-2669)
65
66 QY 34 AlaGlyValIleuValAlaIleuVal 42
67 ||||||||||||||||||||||||
68 DB 998 GCTGAGTGTGTTGTCGCACTTGGTG 1024
69
70 RESULT 6
71 US-10-020-733-3
72 Sequence 3, Application US/10020733

```

Patent No. US20020161214A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Wilgowski, Nathaniel L.
TITLE OF INVENTION: No. US20020161214A1 Human Proteases and Polypeptides
FILE REFERENCE: LEX-0263-USA
CURRENT APPLICATION NUMBER: US/10/020,733
CURRENT FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: US 60/244,939
PRIOR FILING DATE: 2000-11-01
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 2757
TYPE: DNA
ORGANISM: homo sapiens
US-10-020-733-3

Alignment Scores:
Pred. No.: 7 29 Length: 2757
Score: 9.00 Matches: 0
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 77% Indels: 0
DB: 9 Gaps: 0

US-09-831-458A-12 (1-325) x US-10-020-733-3 (1-2757)

Q7 34 AlaclyValLeuValAlaIleLeuVal 42

DB 2113 GCTGAGTCTTCTGCGCATCTTGGTG 2139

RESULT 7

US-10-125-470-2

Sequence 2, Application US/10125470

Patent No. US20020165377A1

GENERAL INFORMATION:

APPLICANT: Popen et al.

TITLE OF INVENTION: ADAM Polynucleotides, polypeptides, and Antibodies

FILE REFERENCE: PPO06P1

CURRENT APPLICATION NUMBER: US/10/125,470

CURRENT FILING DATE: 2002-04-19

PRIOR APPLICATION NUMBER: US/09/712,907A

PRIOR FILING DATE: 2001-03-07

PRIOR APPLICATION NUMBER: 60/136,388

PRIOR FILING DATE: 1999-05-27

PRIOR APPLICATION NUMBER: 60/142,930

PRIOR FILING DATE: 1999-07-09

PRIOR APPLICATION NUMBER: 60/178,717

PRIOR FILING DATE: 2000-01-28

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 2757

TYPE: DNA

ORGANISM: Homo sapiens

US-10-125-470-2

Alignment Scores:
Pred. No.: 7 29 Length: 2757
Score: 9.00 Matches: 0
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 77% Indels: 0
DB: 9 Gaps: 0

US-09-831-458A-12 (1-325) x US-10-125-470-2 (1-2757)

Q7 34 AlaclyValLeuValAlaIleLeuVal 42

DB 2113 GCTGAGTCTTCTGCGCATCTTGGTG 2139

RESULT 8

US-10-125-452-2

Sequence 2, Application US/10125452

Patent No. US20020173640A1

GENERAL INFORMATION:

APPLICANT: Fiedler, Carl Johan

TITLE OF INVENTION: ADAM Polynucleotides, polypeptides, and Antibodies

FILE REFERENCE: PPO06P2

CURRENT APPLICATION NUMBER: US/10/125,452

CURRENT FILING DATE: 2002-04-19

PRIOR APPLICATION NUMBER: 09/745,504

PRIOR FILING DATE: 2001-09-19

PRIOR APPLICATION NUMBER: 09/712,907

PRIOR FILING DATE: 2000-11-16

PRIOR APPLICATION NUMBER: 60/78506/14308

PRIOR FILING DATE: 2000-05-25

PRIOR APPLICATION NUMBER: 60/178,717

PRIOR FILING DATE: 2000-01-28

PRIOR APPLICATION NUMBER: 60/142,930

PRIOR FILING DATE: 1999-07-09

PRIOR APPLICATION NUMBER: 60/136,388

PRIOR FILING DATE: 1999-05-27

NUMBER OF SEQ ID NOS: 48

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 2757

TYPE: DNA

ORGANISM: Homo sapiens

US-10-125-452-2

Alignment Scores:

Pred. No.: 7 29 Length: 2757
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 77% Indels: 0
DB: 9 Gaps: 0

US-09-831-458A-12 (1-325) x US-10-125-452-2 (1-2757)

Q7 34 AlaclyValLeuValAlaIleLeuVal 42

DB 2113 GCTGAGTCTTCTGCGCATCTTGGTG 2139

RESULT 9

US-10-020-733-1

Sequence 1, Application US/10100733

Patent No. US20020161214A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Wilgowski, Nathaniel L.

TITLE OF INVENTION: No. US20020161214A1 Human Proteases and Polypeptides

FILE REFERENCE: LEX-0263-USA

CURRENT APPLICATION NUMBER: US/10/020,733

CURRENT FILING DATE: 2001-10-30

PRIOR APPLICATION NUMBER: 09/744,933

PRIOR FILING DATE: 2000-11-01

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FASTSEQ for Windows Version 4.0

SEQ ID NO 1

LENGTH: 2781

TYPE: DNA

ORGANISM: Homo sapiens

US-10-020-733-1

Alignment Scores:

Pred. No.: 7 35 Length: 2781
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2 77% Indels: 0
DB: 9 Gaps: 0

US-09-831-458A-12 (1-325) x US-10-020-733-1 (1-2781)

QY 34 AlagiyValleuValAla1leuVal 42

DB 2113 GCTGAGTGTGTGTCATCTTGTG 2139

RESULT 10

US-10-020-733-7

Sequence 7, Application US/10020733

Patent No. US20020161214A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Wilganowski, Nathaniel L.

TITLE OF INVENTION: No. US20020161214A1 Human Proteases and Polynucleotides Encodin

FILE REFERENCE: LEX-0263-USA

CURRENT APPLICATION NUMBER: US/10/020,733

PRIOR FILING DATE: 2001-10-30

PRIOR APPLICATION NUMBER: US 60/244,939

PRIOR FILING DATE: 2000-11-01

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 7

LENGTH: 2868

TYPE: DNA

ORGANISM: homo sapiens

US-10-020-733-7

Alignment Scores:

Pred. No.: 7.57

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.77%

DB: 9

Length: 2868

Matches: 9

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-831-458A-12 (1-325) x US-10-020-733-7 (1-2868)

QY 34 AlagiyValleuValAla1leuVal 42

DB 2113 GCTGAGTGTGTGTCATCTTGTG 2139

RESULT 11

US-10-020-733-5

Sequence 5, Application US/10020733

Patent No. US20020161214A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Wilganowski, Nathaniel L.

TITLE OF INVENTION: No. US20020161214A1 Human Proteases and Polynucleotides Encodin

FILE REFERENCE: LEX-0263-USA

CURRENT APPLICATION NUMBER: US/10/020,733

PRIOR FILING DATE: 2001-10-30

PRIOR APPLICATION NUMBER: US 60/244,939

PRIOR FILING DATE: 2000-11-01

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 5

LENGTH: 2892

TYPE: DNA

ORGANISM: homo sapiens

US-10-020-733-5

Alignment Scores:

Pred. No.: 7.64

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.77%

DB: 9

US-09-831-458A-12 (1-325) x US-10-020-733-9 (1-2892)

QY 34 AlagiyValleuValAla1leuVal 42

DB 2113 GCTGAGTGTGTGTCATCTTGTG 2139

RESULT 12

US-10-020-733-9

Sequence 9, Application US/10020733

Patent No. US20020161214A1

GENERAL INFORMATION:

APPLICANT: Walke, D. Wade

APPLICANT: Wilganowski, Nathaniel L.

TITLE OF INVENTION: No. US20020161214A1 Human Proteases and Polynucleotides Encodin

FILE REFERENCE: LEX-0263-USA

CURRENT APPLICATION NUMBER: US/10/020,733

PRIOR FILING DATE: 2001-10-30

PRIOR APPLICATION NUMBER: US 60/244,939

PRIOR FILING DATE: 2000-11-01

NUMBER OF SEQ ID NOS: 9

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO: 9

LENGTH: 3512

TYPE: DNA

ORGANISM: homo sapiens

US-10-020-733-9

Alignment Scores:

Pred. No.: 5.2

Score: 9.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.77%

DB: 9

Length: 3512

Matches: 9

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-831-458A-12 (1-325) x US-10-020-733-9 (1-3512)

QY 34 AlagiyValleuValAla1leuVal 42

DB 2335 GCTGAGTGTGTGTCATCTTGTG 2361

RESULT 13

US-09-894-633A-110/C

Sequence 110, Application US/0984633A

Patent No. US20020124285A1

GENERAL INFORMATION:

APPLICANT: Comer, Timothy

APPLICANT: Dubois, Patrice

APPLICANT: Maiven, Marianne

APPLICANT: Masucci, James

TITLE OF INVENTION: PLANT-BASED ALKALY SUBSTITUTES FOR SELECT

FILE REFERENCE: 38-21(15856)H

CURRENT APPLICATION NUMBER: US/09/894,633A

PRIOR FILING DATE: 2001-06-28

PRIOR APPLICATION NUMBER: 60/214,457

PRIOR FILING DATE: 2000-06-28

PRIOR APPLICATION NUMBER: 09/894,633

PRIOR FILING DATE: 2000-06-28

NUMBER OF SEQ ID NOS: 111

SOFTWARE: PatentIn version 4.0

SEQ ID NO: 110

LENGTH: 623

TYPE: DNA

ORGANISM: Zea mays

US-09-894-633A-110

Alignment Scores:

Pred. No.: 18.7

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 2.46%

DB: 0

Length: 623

Matches: 8

Conservative: 0

Mismatches: 0

Indels: 0

```

DB: 10 Gaps: 0
US-09-831-458a-12 (1-325) x US-09-894-633a-110 (1-625)
Q7 11 gltlenclglenclgtystlen 18
    ||||||||||||||||||
DB 97 CAGCTCGGCTGCGTGGCTGCTG 74

RESULT 14
US-09-770-445-515
: Sequence 515, Application US/09770445
: Patent No. US2002023281A1
: GENERAL INFORMATION:
: APPLICANT: Gorlach, Jorn
: APPLICANT: Au, Yong-Qiang
: APPLICANT: Hamilton, Carol M.
: APPLICANT: Price, Jennifer L.
: APPLICANT: Raines, Tracy M.
: APPLICANT: Yu, Yang
: APPLICANT: Rameaka, Joshua G.
: APPLICANT: Page, Amy
: APPLICANT: Mathew, Abraham V.
: APPLICANT: Ledford, Brooke L.
: APPLICANT: Woessner, Jeffrey P.
: APPLICANT: Haas, William David
: APPLICANT: Garcia, Carlos A.
: APPLICANT: Kricker, Maja
: APPLICANT: Slader, Ted
: APPLICANT: Davis, Keith R.
: APPLICANT: Allen, Keith
: APPLICANT: Holman, Neil
: APPLICANT: Hurlan, Patrick
: TITLE OF INVENTION: Expressed Sequences of Arabidopsis
: FILE REFERENCE: 2023US (PAPA-012PRV)
: CURRENT APPLICATION NUMBER: US/09770,445
: CURRENT FILING DATE: 2001-01-26
: PRIOR FILING DATE: 2000-01-27
: NUMBER OF SEQ ID NOS: 999
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 515
: LENGTH: 884
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-770-445-515

Alignment Scores:
Pred. No.: 26.2 Length: 884
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-770-445-515 (1-884)
Q7 23 lgnValenclnleuSerpho 30
    ||||||||||||||||||
DB 24 CTACTCTCTCACTTTGTGATTT 47

RESULT 15
US-09-894-633a-95/c
: Sequence 95, Application US/09894633A
: Patent No. US20020124285A1
: GENERAL INFORMATION:
: APPLICANT: Conner, Timothy
: APPLICANT: Dubois, Patricia
: APPLICANT: Malvern, Marianne
: APPLICANT: Masucci, James
: TITLE OF INVENTION: PLANT MEMBRANE SEQUENCES FOR SELECTIVE CONFOCAL GENEXPRESSION
: FILE REFERENCE: 38-21(15856)B
: CURRENT APPLICATION NUMBER: US/09/894,633A

```

```

: CURRENT FILING DATE: 2001-06-28
: PRIOR APPLICATION NUMBER: 60/214,357
: PRIOR FILING DATE: 2000-06-28
: PRIOR APPLICATION NUMBER: 09/894,633
: PRIOR FILING DATE: 2000-06-28
: NUMBER OF SEQ ID NOS: 111
: SOFTWARE: PatentIn version 4.0
: SEQ ID NO 95
: LENGTH: 893
: TYPE: DNA
: ORGANISM: Zea mays
US-09-894-633a-95

Alignment Scores:
Pred. No.: 26.4 Length: 893
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.46% Indels: 0
DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-894-633a-95 (1-893)
Q7 11 gltlenclglenclgtystlen 18
    ||||||||||||||||||
DB 367 CAGCTCGGCTGCGTGGCTGCTG 344

RESULT 16
US-09-770-445-328/c
: Sequence 328, Application US/09770445
: Patent No. US2002023281A1
: GENERAL INFORMATION:
: APPLICANT: Gorlach, Jorn
: APPLICANT: Au, Yong-Qiang
: APPLICANT: Hamilton, Carol M.
: APPLICANT: Price, Jennifer L.
: APPLICANT: Raines, Tracy M.
: APPLICANT: Yu, Yang
: APPLICANT: Rameaka, Joshua G.
: APPLICANT: Page, Amy
: APPLICANT: Mathew, Abraham V.
: APPLICANT: Ledford, Brooke L.
: APPLICANT: Woessner, Jeffrey P.
: APPLICANT: Haas, William David
: APPLICANT: Garcia, Carlos A.
: APPLICANT: Kricker, Maja
: APPLICANT: Slader, Ted
: APPLICANT: Davis, Keith R.
: APPLICANT: Allen, Keith
: APPLICANT: Holman, Neil
: APPLICANT: Hurlan, Patrick
: TITLE OF INVENTION: Expressed Sequences of Arabidopsis
: FILE REFERENCE: 2023US (PAPA-012PRV)
: CURRENT APPLICATION NUMBER: US/09770,445
: CURRENT FILING DATE: 2001-01-26
: PRIOR FILING DATE: 2000-01-27
: NUMBER OF SEQ ID NOS: 999
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 328
: LENGTH: 952
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(952)
: OTHER INFORMATION: n = A,T,C or G
US-09-770-445-328

Alignment Scores:
Pred. No.: 28.1 Length: 952
Score: 8.00 Matches: 8

```

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.46% Indels: 0
 DB: 10 Gaps: 0

US-09-831-458A-12 (1-325) x US-09-770-445-328 (1-522)

UY 262 CluGlyThrTp6InTPvAlasp 269

DB 712 CAAGGACCTTGCAATGCGTTCAT 689

RESULT 17

US-09-887-576-673

Sequence 673, Application US/09887576
 Patent No. US20020144047A1

GENERAL INFORMATION:

APPLICANT: Budworth, P.

APPLICANT: Brown, D.

APPLICANT: Chang, H.

APPLICANT: Zhu, T.

APPLICANT: Han, B.

APPLICANT: Wang, X.

APPLICANT: Cooper, Bret

FILE OF INVENTION: Promoters for regulation of plant expression

TITLE REFERENCE: 1360.001US1

CURRENT APPLICATION NUMBER: US/09/887,576

PRIOR FILING DATE: 2001-06-25

PRIOR APPLICATION NUMBER: US 60/213,848

PRIOR FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: US 60/214,087

PRIOR FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: US 60/258,692

PRIOR FILING DATE: 2000-12-29

NUMBER OF SEQ ID NOS: 875

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 673

LENGTH: 1149

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-887-576-673

Alignment Scores:
 Pred. No.: 33.7 Length: 1149
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.46% Indels: 0
 DB: 10 Gaps: 0

US-09-831-458A-12 (1-325) x US-09-887-576-673 (1-1149)

UY 47 ValProSerLeuSerGlnGlu 54

DB 493 GTGGACATCTGTTCTTCATGAG 516

RESULT 18

US-09-887-576-826

Sequence 826, Application US/09887576
 Patent No. US20020144047A1

GENERAL INFORMATION:

APPLICANT: Budworth, P.

APPLICANT: Brown, D.

APPLICANT: Chang, H.

APPLICANT: Zhu, T.

APPLICANT: Han, B.

APPLICANT: Wang, X.

APPLICANT: Cooper, Bret

FILE OF INVENTION: Promoters for regulation of plant expression

TITLE REFERENCE: 1360.001US1

CURRENT APPLICATION NUMBER: US/09/887,576

PRIOR FILING DATE: 2001-06-25

PRIOR APPLICATION NUMBER: US 60/213,848

PRIOR FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: US 60/213,848
 PRIOR FILING DATE: 2000-06-23
 PRIOR APPLICATION NUMBER: US 60/258,692
 PRIOR FILING DATE: 2000-12-29
 NUMBER OF SEQ ID NOS: 875
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 826

US-09-831-458A-12 (1-325) x US-09-887-576-826 (1-2000)

UY 12 LeuGlyLeuLeuGlyGlyGlyGly 19

DB 1355 TTAGCTCTATGATGTTGTTTACAG 1478

RESULT 19

US-09-962-436-341/c

Sequence 331, Application US/09962436
 Patent No. US20020081301A1

GENERAL INFORMATION:

APPLICANT: Sappel, Daniel

FILE OF INVENTION: Cancer Gene Determination and Therapy

TITLE OF INVENTION: Sets

FILE REFERENCE: 689290-75

CURRENT APPLICATION NUMBER: US/09/962,436

CURRENT FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US/09/962,436

PRIOR FILING DATE: 2000-09-25

PRIOR APPLICATION NUMBER: US/09/962,436

PRIOR FILING DATE: 2000-09-25

PRIOR APPLICATION NUMBER: US/09/962,436

PRIOR FILING DATE: 2000-09-25

NUMBER OF SEQ ID NOS: 568

SOFTWARE: Patent In version 4.0

SEQ ID NO 331

LENGTH: 3150

TYPE: DNA

ORGANISM: Homo sapiens

US-09-962-436-331

Alignment Scores:
 Pred. No.: 89 Length: 3150
 Score: 8.00 Matches: 3
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.46% Indels: 0
 DB: 10 Gaps: 0

US-09-831-458A-12 (1-325) x US-09-962-436-331 (1-3150)

UY 38 ValAlaIleLeuValGlnValSer 45

DB 1085 GTAGCATCTGTTGTTGTTGTTGTT 1066

RESULT 20

US-09-969-708-453/c

Sequence 453, Application US/09969708
 Patent No. US20020102532A1

GENERAL INFORMATION:

APPLICANT: Augustus, Meena

FILE OF INVENTION: Cancer Gene Determination and Therapy

TITLE OF INVENTION: Sets

FILE REFERENCE: 689290-70

CURRENT APPLICATION NUMBER: US/09/969,708

PRIOR FILING DATE: 2000-06-23

PRIOR APPLICATION NUMBER: US/09/969,708

PRIOR FILING DATE: 2000-06-23

```

CURRENT FILING DATE: 2001-10-03
PRIOR APPLICATION NUMBER: US/60/237,606
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237,608
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237,425
PRIOR FILING DATE: 2000-10-03
NUMBER OF SEQ ID NOS: 658
SOFTWARE: PatentIn version 3.0
SEQ ID NO 453
LENGTH: 3150
TYPE: DNA
ORGANISM: Homo sapiens
US-09-969-708-453

Alignment Scores:
Pred. No.: 89
Score: 8.00
Length: 3150
Percent similarity: 100.00%
Matches: 8
Best local similarity: 100.00%
Conservative: 0
Mismatch: 0
Query Match: 2.46%
Indels: 0
DB: 10

US-09-831-458a-12 (1-325) x US-09-969-708-453 (1-3150)
CY 38 ValAla11eValGlnValSer 45
|||||
DB 1089 GTACCTATCTTCGTCACGCTCTCC 1086

RESULT 21
US-09-954-456-1224/c
Sequence 1224, Application US/99/44456
Patent No. US2001015057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
FILE REFERENCE: 689290-76
CURRENT APPLICATION NUMBER: US/00/954,456
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/00/924,200
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,127
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1224
LENGTH: 3150
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1224

Alignment Scores:
Pred. No.: 89
Score: 8.00
Length: 3150
Percent similarity: 100.00%
Matches: 8
Best local similarity: 100.00%
Conservative: 0
Mismatch: 0
Query Match: 2.46%
Indels: 0
DB: 10

```

```

Query Match: 2.46%
Indels: 0
DB: 10

US-09-831-458a-12 (1-325) x US-09-954-456-1224 (1-3150)
CY 38 ValAla11eValGlnValSer 45
|||||
DB 1089 GTACCTATCTTCGTCACGCTCTCC 1086

RESULT 22
US-09-764-847-1317/c
Sequence 1317, Application US/97/64847
Patent No. US20020132767A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: 67004
CURRENT APPLICATION NUMBER: US/97/764,847
CURRENT FILING DATE: 2001-01-17
PRIOR APPLICATION data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 2004
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1317
LENGTH: 18871
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-847-1317

Alignment Scores:
Pred. No.: 499
Score: 8.00
Length: 18871
Percent similarity: 100.00%
Matches: 8
Best local similarity: 100.00%
Conservative: 0
Mismatch: 0
Query Match: 2.46%
Indels: 0
DB: 10

US-09-831-458a-12 (1-325) x US-09-764-847-1317 (1-18871)
CY 44 ValSer15ValProGSerGSer 51
|||||
DB 1984 GTTCGACGCTTCATCCAGCTTG 1961

RESULT 23
US-09-294-093B-2487/c
Sequence 2487, Application US/97/4093B
Patent No. US20010051335A1
GENERAL INFORMATION:
APPLICANT: Ito, Taisugi, Kashimatsu, Y.
APPLICANT: Sherman, Bradley, N.
TITLE OF INVENTION: POLYPEPTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
FILE REFERENCE: PL-0009 US
CURRENT APPLICATION NUMBER: US/97/294,093B
CURRENT FILING DATE: 1999-04-16
PRIOR APPLICATION NUMBER: US/97/242,567
PRIOR FILING DATE: April 21, 1998
NUMBER OF SEQ ID NOS: 6207
SOFTWARE: PERL Program
SEQ ID NO 2487
LENGTH: 124
TYPE: DNA
ORGANISM: Zoa mays
FEATURE:
NAME/KEY: misc. feature
OTHER INFORMATION: Incyte ID No. US20010051335A1 700346184H1
NAME/KEY: unsure
LOCATION: 5', 40, 69, 73
OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-2487

Alignment Scores:
Pred. No.: 42.4
Score: 7.00
Length: 124
Matches: 7

```


Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
DB: 10
Gaps: 0

US-09-831-458a-12 (1-325) x US-09-294-0938 2487 (1-124)

UY 45 GYVAlleuValAlaIleleu 41
|||||

Db 21 GCGCTGCTGTGCAATACCTC 1

RESULT 24
US-10-046-935-1658
Sequence 1658, Application US/10046935
Patent No. US2002015601A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yujin
APPLICANT: Harlocker, Susan L.
APPLICANT: Secrist, Heather
APPLICANT: Wang, Aljun
APPLICANT: Stork, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.52721
CURRENT APPLICATION NUMBER: US/10/046,935
CURRENT FILING DATE: 2002-01-15
NUMBER OF SEQ. ID NOS: 2239
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1658
LENGTH: 128
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9, 10, 11, 53, 68
OTHER INFORMATION: n = A,T,C or G
US-10-046-935-1658

Alignment Scores:
Pred. No.: 43.7
Score: 7.00
Length: 128
Matches: 7
Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
Indels: 0
DB: 9
Gaps: 0

US-09-831-458a-12 (1-325) x US-10-046-935-1658 (1-128)

UY 10 GlnGlnLeuGlyLeuLeuGly 16
|||||

Db 104 CAGCAACTGCGCTTACTAGCG 124

RESULT 25
US-09-878-178-1658
Sequence 1658, Application US/09878178
Patent No. US20020177552A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yujin
APPLICANT: Harlocker, Susan L.
APPLICANT: Secrist, Heather
APPLICANT: Wang, Aljun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.527
CURRENT APPLICATION NUMBER: US/09/878,178
CURRENT FILING DATE: 2001-06-08
NUMBER OF SEQ. ID NOS: 2237
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1658
LENGTH: 128
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature

LOCATION: (1)-(128)
OTHER INFORMATION: n = A,T,C or G
US-09-878-178-1658

Alignment Scores:
Pred. No.: 43.7
Score: 7.00
Length: 128
Matches: 7
Conservative: 0
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 2.15%
Indels: 0
DB: 9
Gaps: 0

US-09-831-458a-12 (1-325) x US-09-878-178-1658 (1-128)

UY 10 GlnGlnLeuGlyLeuLeuGly 16
|||||

Db 104 CAGCAACTGCGCTTACTAGCG 124

RESULT 26
US-09-864-761-23399/c
Sequence 23399, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Rank, David R.
APPLICANT: Hauzel, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME DERIVED SINGLE EXON NUCLE
FILE REFERENCE: Acomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,412
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/654,2,466
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GR 242644
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: n: 60/246,459
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/178501/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/178501/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/178501/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/178501/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/178501/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/178501/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/178501/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/178501/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/178501/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/178501/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/244,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/6108,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ. ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Editing vers. 1.1
SEQ ID NO 23399
LENGTH: 238
TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO A121586.28
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.8
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.3
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.6
 OTHER INFORMATION: EXPRESSED IN HEPA2, SIGNAL = 2.1
 OTHER INFORMATION: EXPRESSED IN HEPA1, SIGNAL = 2.2
 OTHER INFORMATION: NOT HIT: A114254.1, PVALUE 2.20e+00
 OTHER INFORMATION: SWISSPROT HIT: Q11205, EVALUATE 1.20e+02
 US-09-864-761-23399

Alignment Scores:
 Pred. No.: 79.5 Length: 238
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.15% Indels: 0
 DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-864-761 23399 (1-238)

Q7 32 IGVLEUVALAGLYVALGVYAL 38
 |||||
 DL 36 CTCTGGCTGGAGTCTGTGATG 16

RESULT 27
 US-09-923-876-704

Sequence 704, Application US/09023876
 Patent No. US20020013958A1
 GENERAL INFORMATION:
 APPLICANT: Kamiyaki, Laura Y. (Ito)
 APPLICANT: Sherman, Bradley K.
 TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
 FILE REFERENCE: PL-0012-1 CON
 CURRENT APPLICATION NUMBER: US/09/923,876
 CURRENT FILING DATE: 2001-08-06
 PRIOR APPLICATION NUMBER: 09/294,429
 PRIOR FILING DATE: 1999-04-21
 PRIOR APPLICATION NUMBER: 60/085,331
 PRIOR FILING DATE: 1998-05-05
 NUMBER OF SEQ ID NOS: 6332
 SOFTWARE: PERL Program
 SEQ ID NO 704
 LENGTH: 258
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Incyte ID No. US20020013958A1 700157413H
 NAME/KEY: unsure
 LOCATION: 67..112
 OTHER INFORMATION: a, t, c, g, or other
 US-09-923-876-704

Alignment Scores:
 Pred. No.: 85.9 Length: 258
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.15% Indels: 0
 DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-923-876-704 (1-258)

Q7 285 GIVPQASNASISERGLYSAN 291
 |||||
 DL 14 GAGCTTAACCAACAGTCGGAAC 34

RESULT 28
 US-09-974-300-3994/6

Sequence 3994, Application US/09974300
 Patent No. US20020145721A1
 GENERAL INFORMATION:
 APPLICANT: Petka, Randy M.
 APPLICANT: Clauson, ID Groll
 TITLE OF INVENTION: Methods for Monitoring Multiple Gene
 TITLE OF INVENTION: Expression
 FILE REFERENCE: 10085,509-05
 CURRENT APPLICATION NUMBER: US/09/974,300
 CURRENT FILING DATE: 2001-10-05
 PRIOR APPLICATION NUMBER: 09/680,598
 PRIOR FILING DATE: 2000-10-06
 PRIOR APPLICATION NUMBER: 60/279,526
 PRIOR FILING DATE: 2001-03-27
 NUMBER OF SEQ ID NOS: 8481
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3994
 LENGTH: 258
 TYPE: DNA
 ORGANISM: Bacillus licheniformis
 US-09-974-300-3994

Alignment Scores:
 Pred. No.: 85.9 Length: 258
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.15% Indels: 0
 DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-974-300-3994 (1-258)

Q7 112 TGVAGLVGVVALGVYAL 118
 |||||
 DL 173 ACCGCTTAACCAACGCGAGTC 153

RESULT 29

US-09-923-876-4825/6
 Sequence 4825, Application US/09923876
 Patent No. US20020013958A1
 GENERAL INFORMATION:
 APPLICANT: Kamiyaki, Laura Y. (Ito)
 APPLICANT: Sherman, Bradley K.
 TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
 FILE REFERENCE: PL-0012-1 CON
 CURRENT APPLICATION NUMBER: US/09/923,876
 CURRENT FILING DATE: 2001-08-06
 PRIOR APPLICATION NUMBER: 09/294,429
 PRIOR FILING DATE: 1999-04-21
 PRIOR APPLICATION NUMBER: 60/085,331
 PRIOR FILING DATE: 1998-05-05
 NUMBER OF SEQ ID NOS: 6332
 SOFTWARE: PERL Program
 SEQ ID NO 4825
 LENGTH: 258
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Incyte ID No. US20020013958A1 700455864H1
 NAME/KEY: unsure
 LOCATION: 45..186, 222
 OTHER INFORMATION: a, t, c, g, or other
 US-09-923-876-4825

Alignment Scores:
 Pred. No.: 89.1 Length: 268
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-923-876-4825 (1-268)

QY 35 G1yValLeuValAlaIleLeu 41
DB 123 GCGGTGCTTTCATCATCTC 103

RESULT 30

US-09-923-876-4499/c

Sequence 4499, Application US/09924876
Patent No. US20020013958A1

GENERAL INFORMATION:

APPLICANT: Laliqadi, Raghunath V.

APPLICANT: Kamigaki, Laura Y. (ILO)

APPLICANT: Sherman, Bradley K.

TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING

FILE REFERENCE: PI-0012-1-COR

CURRENT APPLICATION NUMBER: US/09/923,876

CURRENT FILING DATE: 2001-08-06

PRIOR APPLICATION NUMBER: 09/298,329

PRIOR FILING DATE: 1999-04-21

PRIOR APPLICATION NUMBER: 60/085,331

PRIOR FILING DATE: 1998-05-05

NUMBER OF SEQ. ID NOS: 4332

SOFTWARE: PERL Program

SEQ. ID NO: 4499

LENGTH: 275

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: Intron ID No. US20020013958A1 790455143H1

LOCATION: 257

OTHER INFORMATION: a, t, c, g, or other

US-09-923-876-4499

Alignment Scores:
Pred. No.: 91.3 Length: 275
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-923-876-4499 (1-275)

QY 35 G1yValLeuValAlaIleLeu 41
DB 127 GCGGTGCTTTCATCATCTC 107

RESULT 31

US-09-878-574-13627/c

Sequence 13627, Application US/09878574
Patent No. US20020110548A1

GENERAL INFORMATION:

APPLICANT: Byrum, Joseph R.

APPLICANT: La Kosa, Thomas J.

APPLICANT: Thompson, Michael D.

TITLE OF INVENTION: Nucleic Acid Molecules and other Molecules Associated with

FILE REFERENCE: 38-21(15401)B

CURRENT APPLICATION NUMBER: US/09/878,574

CURRENT FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 09/333,535

PRIOR FILING DATE: 1999-06-14

NUMBER OF SEQ. ID NOS: 15775

SEQ. ID NO: 13627

LENGTH: 275

TYPE: DNA

ORGANISM: Glycine max

OTHER INFORMATION: Clone ID: 701067461H1
US-09-878-574-13627

Alignment Scores:

Pred. No.: 91.3 Length: 275
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-878-574-13627 (1-275)

QY 271 SerProLeuSerProSerPhe 277
DB 21 TCGGCACCTTCACCATCTTC 1

RESULT 32

US-09-764-869-2051

Sequence 2051, Application US/09764869
Patent No. US20020061521A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC007

CURRENT APPLICATION NUMBER: US/09/764,869

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file with

NUMBER OF SEQ. ID NOS: 2442

SOFTWARE: PatentIn Ver. 2.0

SEQ. ID NO: 2051

LENGTH: 314

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-869-2051

Alignment Scores:
Pred. No.: 104 Length: 314
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0

US-09-831-458a-12 (1-325) x US-09-764-869-2051 (1-314)

QY 27 LeuLeuSerPheMetLeuLeu 33
DB 126 CTCCTTCTTCATGTTGCTG 146

RESULT 33

US-09-764-869-2052

Sequence 2052, Application US/09764869
Patent No. US20020061521A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC007

CURRENT APPLICATION NUMBER: US/09/764,869

CURRENT FILING DATE: 2001-01-17

Prior application data removed - refer to PALM or file with

NUMBER OF SEQ. ID NOS: 2442

SOFTWARE: PatentIn Ver. 2.0

SEQ. ID NO: 2052

LENGTH: 314

TYPE: DNA

ORGANISM: Homo sapiens

US-09-764-869-2052

Alignment Scores:
Pred. No.: 104 Length: 314
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0

US-09-831-458A-12 (1-325) x US-09-764-869-2052 (1-314)
QY 27 LeuleuserPheMeLeuserPhe 33
|||||
DB 126 CIGCTTCTTCATGTTGCTG 146

RESULT 34
US-09-983-965-1896
: Sequence 1896, Application US/0909465
: Patent No. US20020137160A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Ningbing
: APPLICANT: Hyatt, John C.
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LATENTATION AND
: FILE REFERENCE: 37-21(10297)C
: CURRENT APPLICATION NUMBER: US/09/093,965
: CURRENT FILING DATE: 2001-10-26
: PRIOR APPLICATION NUMBER: US 09/465,231
: PRIOR FILING DATE: 1999-12-15
: PRIOR APPLICATION NUMBER: US 60/113,678
: PRIOR FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 5912
: SEQ ID NO 1896
: LENGTH: 326
: TYPE: DNA
: ORGANISM: Bos taurus
: FEATURE:
: OTHER INFORMATION: Clone ID: 11-L1B3057-007-Q1-K1-C3
US-09-983-965-1896

Alignment Scores:
Pred. No.: 108 Length: 326
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0

US-09-831-458A-12 (1-325) x US-09-983-965-1896 (1-326)
QY 24 ValLeuGInLeuLeuserPhe 30
|||||
DB 106 CTTTAAAGCTCTGTGAATTT 126

RESULT 35
US-09-974-300-2011/C
: Sequence 2011, Application US/09974300
: Patent No. US20020146721A1
: GENERAL INFORMATION:
: APPLICANT: Berka, Randy M.
: APPLICANT: Clausen, Ib Groth
: TITLE OF INVENTION: Methods For Monitoring Multiple Gene
: FILE REFERENCE: 10085, 500-US
: CURRENT APPLICATION NUMBER: US/09/974,300
: CURRENT FILING DATE: 2001-10-05
: PRIOR APPLICATION NUMBER: 09/640,548
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: 60/279,526
: PRIOR FILING DATE: 2001-03-27
: NUMBER OF SEQ ID NOS: 8481
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 2011
: LENGTH: 346
: TYPE: DNA
: ORGANISM: Sacillus licheniformis

```

```

US-09-974-300-2011
Alignment Scores:
Pred. No.: 111 Length: 336
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0

US-09-831-458A-12 (1-325) x US-09-974-300-2011 (1-336)
QY 23 LeCValLeuGInLeuLeuser 29
|||||
DB 22 TTAACTTGTGACCTCTTCA 2

RESULT 36
US-09-810-997-8/C
: Sequence 8, Application US/09810997
: Patent No. US2002009501A1
: GENERAL INFORMATION:
: APPLICANT: Song, Xiaoling
: APPLICANT: Fan, Hao
: APPLICANT: Wei, Zhong-Min
: TITLE OF INVENTION: RECEPTORS FOR HYPERSENSITIVE RESPONSE ELICITORS AND
: FILE REFERENCE: 21829/62
: CURRENT APPLICATION NUMBER: US/09/810,997
: CURRENT FILING DATE: 2001-03-16
: PRIOR APPLICATION NUMBER: 60/191,649
: PRIOR FILING DATE: 2000-03-23
: PRIOR APPLICATION NUMBER: 60/250,710
: PRIOR FILING DATE: 2000-12-01
: NUMBER OF SEQ ID NOS: 9
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 8
: LENGTH: 342
: TYPE: DNA
: ORGANISM: Xanthomonas campestris pv. pelargonii
US-09-810-997-8

Alignment Scores:
Pred. No.: 113 Length: 342
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15% Indels: 0
DB: 10 Gaps: 0

US-09-831-458A-12 (1-325) x US-09-810-997-8 (1-342)
QY 32 LeuLeuAlaGlyValLeuVal 38
|||||
DB 90 CTGTGTGAGTGGAGTCTCTGTA 70

RESULT 37
US-09-880-371-14/C
: Sequence 14, Application US/09880371
: Patent No. US20020059658A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: APPLICANT: DeRoche, Jay
: TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
: FILE REFERENCE: 21829/91
: CURRENT APPLICATION NUMBER: US/09/880,371
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: 60/211,585
: PRIOR FILING DATE: 2000-06-15
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 14
: LENGTH: 342

```

```

: TYPE: DNA
: ORGANISM: Xanthomonas campestris
US-09-880-371-14

Alignment Scores:
Pred. No.: 113          Length: 342
Score: 7.00            Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.15%              Indels: 0
DB: 10                      Gaps: 0

US-09-831-458A-12 (1-325) x US-09-880-371-14 (1-342)

QY 32 LeuleuAlaGlyValLeuVal 38
DB 90 CTGCTGCTGAGTCTGCTG 70

RESULT 38
US-09-880-371-15/C
: Sequence 15, Application US/09880371
: Patent No. US20020059658A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: APPLICANT: Derocher, Jay
: TITLE OF INVENTION: METHODS OF IMPROVING THE EFFECTIVENESS OF TRANSGENIC
: FILE REFERENCE: 21829/91
: CURRENT APPLICATION NUMBER: US/09/880,371
: CURRENT FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: 60/211,585
: PRIOR FILING DATE: 2000-06-15
: NUMBER OF SEQ ID NOS: 16
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO 15
: LENGTH: 342
: TYPE: DNA
: ORGANISM: Nicotiana tabacum
US-09-880-371-15

Alignment Scores:
Pred. No.: 113          Length: 342
Score: 7.00            Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.15%              Indels: 0
DB: 10                      Gaps: 0

US-09-831-458A-12 (1-325) x US-09-880-371-15 (1-342)

QY 32 LeuleuAlaGlyValLeuVal 38
DB 90 CTGCTGCTGAGTCTGCTG 70

RESULT 39
US-09-829-124-1/C
: Sequence 1, Application US/09829124
: Patent No. US20020066122A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: APPLICANT: Swanson, Shane S.
: TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM XANTHOMAS
: FILE REFERENCE: 21829/101
: CURRENT APPLICATION NUMBER: US/09/829,124
: CURRENT FILING DATE: 2001-04-09
: PRIOR APPLICATION NUMBER: 60/224,053
: PRIOR FILING DATE: 2000-08-09
: PRIOR APPLICATION NUMBER: 09/412,452
: PRIOR FILING DATE: 1999-10-04
: PRIOR APPLICATION NUMBER: 60/103,124
: PRIOR FILING DATE: 1998-10-05

```

```

: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: Patent Ver. 2.1
: SEQ ID NO 1
: LENGTH: 342
: TYPE: DNA
: ORGANISM: Xanthomonas campestris
US-09-829-124-1

Alignment Scores:
Pred. No.: 113          Length: 342
Score: 7.00            Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.15%              Indels: 0
DB: 10                      Gaps: 0

US-09-831-458A-12 (1-325) x US-09-829-124-1 (1-342)

QY 32 LeuleuAlaGlyValLeuVal 38
DB 90 CTGCTGCTGAGTCTGCTG 70

RESULT 40
US-09-983-965-2185
: Sequence 2185, Application US/09983965
: Patent No. US20020137160A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nengping
: APPLICANT: Byatt, John C.
: APPLICANT: Mathiadagu, Nadappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOC
: FILE REFERENCE: 37-21(110297)C
: CURRENT APPLICATION NUMBER: US/09/983,965
: CURRENT FILING DATE: 2001-10-26
: PRIOR APPLICATION NUMBER: US-09/465,231
: PRIOR FILING DATE: 1999-12-15
: PRIOR APPLICATION NUMBER: US 60/114,678
: PRIOR FILING DATE: 1998-12-17
: NUMBER OF SEQ ID NOS: 5912
: SEQ ID NO 2185
: LENGTH: 349
: TYPE: DNA
: ORGANISM: Bos taurus
: FEATURE:
: OTHER INFORMATION: Clone ID: 43-1184057-007-Q1-K1-C4
US-09-983-965-2185

Alignment Scores:
Pred. No.: 115          Length: 349
Score: 7.00            Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 2.15%              Indels: 0
DB: 10                      Gaps: 0

US-09-831-458A-12 (1-325) x US-09-983-965-2185 (1-349)

QY 24 ValLeuGlnLeuLeuSerPhe 40
DB 168 GTTTACAGCTGCTGAGTTT 168

RESULT 41
US-09-867-701-3187
: Sequence 3187, Application US/09867701
: Patent No. US2002013237A1
: GENERAL INFORMATION:
: APPLICANT: Aglate, Paul A.
: APPLICANT: Jones, Robert
: APPLICANT: Harlocker, Susan L.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: OTHER INFORMATION: AND DIAGNOSIS OF OVARIAN CANCER

```

```

: FILE REFERENCE: 210121.497
: CURRENT APPLICATION NUMBER: US/09/867,701
: CURRENT FILING DATE: 2001-05-29
: NUMBER OF SEQ ID NOS: 10912
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO: 3187
: LENGTH: 454
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-867-701-3187

Alignment Scores:
Pred. No.: 117          Length: 354
Score: 7.00            Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15%      Indels: 0
DB: 10                  Gaps: 0

US-09-831-458a-12 (1-325) x US-09-867-701-3187 (1-354)
QY 243 10cgccthrseratserasu 249
DB 333 TCGCAAGACTCTAGAGCAAT 353

RESULT 42
US-09-880-107-1587
: Sequence 158, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scheff, Uwe
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921.5028.WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO: 158
: LENGTH: 363
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 AAC71387
US-09-880-107-158

Alignment Scores:
Pred. No.: 119          Length: 363
Score: 7.00            Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15%      Indels: 0
DB: 10                  Gaps: 0

US-09-831-458a-12 (1-325) x US-09-880-107-158 (1-363)
QY 4 Setlysgluhrarqvalslu 10
DB 433 TCCAAAGACCAAGACTGCCA 313

RESULT 43
US-09-878-574-1827
: Sequence 182, Application US/09878574
: Patent No. US20020119548A1
: GENERAL INFORMATION:
: APPLICANT: Byrum, Joseph R.
: APPLICANT: La Rosa, Thomas J.

```

```

: APPLICANT: Thompson, Michael D.
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: FILE REFERENCE: 38-21(15401)B
: CURRENT APPLICATION NUMBER: US/09/878,574
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: 09/434,535
: PRIOR FILING DATE: 1999-06-14
: NUMBER OF SEQ ID NOS: 15775
: SEQ ID NO: 182
: LENGTH: 364
: TYPE: DNA
: ORGANISM: Glycine max
: OTHER INFORMATION: Clone ID: LIB3028-054-Q1-B1-D6
US-09-878-574-182

Alignment Scores:
Pred. No.: 120          Length: 364
Score: 7.00            Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15%      Indels: 0
DB: 10                  Gaps: 0

US-09-831-458a-12 (1-325) x US-09-878-574-182 (1-364)
QY 271 scfrrtscrrtscrrtscrrt 277
DB 73 TCGCTACTTTCACCATCTTTC 53

RESULT 44
US-09-960-352-2949
: Sequence 2949, Application US/09960352
: Patent No. US20020147139A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Mengding
: APPLICANT: Byatt, John C.
: APPLICANT: Mathalaagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
: FILE REFERENCE: 16511.036/27-21(10298)C
: CURRENT APPLICATION NUMBER: US/09/960,352
: CURRENT FILING DATE: 2001-09-24
: NUMBER OF SEQ ID NOS: 15112
: SEQ ID NO: 2949
: LENGTH: 384
: TYPE: DNA
: ORGANISM: Bos taurus
: OTHER INFORMATION: Clone ID: 13 LTB34-607-Q1-E1-D1
US-09-960-352-2949

Alignment Scores:
Pred. No.: 126          Length: 484
Score: 7.00            Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.15%      Indels: 0
DB: 10                  Gaps: 0

US-09-831-458a-12 (1-325) x US-09-960-352-2949 (1-384)
QY 269 Aspclyscrrtscrrtscrrt 275
DB 23 GATCGAACCCCTCTGTCACCA 43

RESULT 45
US-09-826-124-677
: Sequence 6, Application US/09826124
: Patent No. US20020066122A1
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: APPLICANT: Swanson, Shane S.

```


•
•
•
•

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd

OM protein - protein search, using sw model

Run on: December 7, 2002 11:00:15 Search time: 49 Seconds
(without alignments)
1562.184 Million cell updates/sec

Title: US-09-831-458A-12

Percent score: 325

Sequence: 1 MSLSKPRVQQLNLGLSLGH PDDVNYWICKKPAAPRDE 325

Scoring table: OLIGO Gapop 60.0, Capext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 5

Total number of hits satisfying chosen parameters: 17736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database: PIP-73.3
1: p1r1:
2: p1r2:
3: p1r3:
4: p1r4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	24.0	404	2 A46274	HIV gp120-binding
2	9	2.8	168	2 B88102	protein W09G10.5 (
3	8	2.5	214	2 E82350	orotate phosphorib
4	8	2.5	279	2 T52454	ATP-dependent Clp
5	8	2.5	282	2 S74644	regulatory compone
6	8	2.5	382	2 T01943	hypothetical prote
7	7	2.2	129	1 DEECS4	succinate dehydrog
8	7	2.2	129	2 H90722	succinate dehydrog
9	7	2.2	129	2 A85573	succinate dehydrog
10	7	2.2	139	2 AH0590	succinate dehydrog
11	7	2.2	130	2 A82119	succinate dehydrog
12	7	2.2	160	2 B70302	hypothetical prote
13	7	2.2	161	2 S76399	hypothetical prote
14	7	2.2	207	2 T21881	hypothetical prote
15	7	2.2	210	2 S43399	GTP-binding protei
16	7	2.2	227	2 T35281	probable two compo
17	7	2.2	232	2 D85654	probable uracil ac
18	7	2.2	235	2 AF3451	1198 protein (limp
19	7	2.2	239	2 F71343	hypothetical prote
20	7	2.2	240	2 S65218	hypothetical prote
21	7	2.2	244	2 T26913	hypothetical prote
22	7	2.2	244	2 T26912	hypothetical prote
23	7	2.2	246	2 P95397	hypothetical prote
24	7	2.2	247	2 A99794	Probable halocid
25	7	2.2	247	2 G82956	uracil associated
26	7	2.2	275	2 A64145	hypothetical prote
27	7	2.2	275	2 C81843	hypothetical prote
28	7	2.2	275	2 C81097	cutoxylase (
29	7	2.2	275	2 A97226	esterase, probable
30	7	2.2	275	2 A97226	uncharacterized co

30	7	2.2	283	2 T06278	Protein W09G10.5 (
31	7	2.2	288	2 B97074	Protein W09G10.5 (
32	7	2.2	296	2 A22847	Protein W09G10.5 (
33	7	2.2	304	2 T30209	Protein W09G10.5 (
34	7	2.2	304	2 H81449	Protein W09G10.5 (
35	7	2.2	312	2 E37614	Protein W09G10.5 (
36	7	2.2	338	2 E71957	Protein W09G10.5 (
37	7	2.2	338	2 E64521	Protein W09G10.5 (
38	7	2.2	352	2 T05723	Protein W09G10.5 (
39	7	2.2	367	2 F64847	Protein W09G10.5 (
40	7	2.2	367	2 A85624	Protein W09G10.5 (
41	7	2.2	367	2 C90760	Protein W09G10.5 (
42	7	2.2	368	2 D64101	Protein W09G10.5 (
43	7	2.2	376	2 A95447	Protein W09G10.5 (
44	7	2.2	392	2 D96770	Protein W09G10.5 (
45	7	2.2	394	2 D86908	Protein W09G10.5 (

ALIGNMENTS

RESULT 1

A46274

HIV gp120-binding C-type lectin - human

C:Species: Homo sapiens (man)

C>Date: 21-Sep-1993 *Sequence-Revision: 18 Nov 1994 *Level: cont

C:Accession: A46274

R:Curtris, B.M.; Scharnowski, S.; Watson, A.J.

Proc. Natl Acad. Sci. U S A 89, 8456-8460, 1992

A:Title: Sequence and expression of a membrane associated C-

A:Reference number: A46274; M01092304446; PMID:1518869

A:Accession: A46274

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-404 <CUR>

A:Experimental source: placenta

A:Note: sequence extracted from NCBI backbone (NCBI:113134)

C:Superfamily: C-type lectin homology

F:256-377/domain: C-type lectin homology - LCH.

Query Match

Best Local Similarity 100.0%; Score 78; DB 2; Length 4

Matches 78; Conservative 0; Mismatches 0; Indel

QY	58	QATVNTIQLKAAGSELSKSLGELLYGLKAAVGLPKSKLQET	17
hb	74	QATVNTIQLKAAGSELSKSLGELLYGLKAAVGLPKSKLQET	17
QY	118	VGELPKSKLQELIYQELT 145	17
hb	134	VGELPKSKLQELIYQELT 151	17

RESULT 2

B88102

protein W09G10.5 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 *Sequence-Revision: 10 May 2001 *Text: ch

C:Accession: B88102

C:Anonymous: The C. elegans Sequencing Consortium.

Science 287, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a plat

A:Reference number: A75000; M01092304446; PMID:9851916

A:Note: seq webstis genome webst. 04/09/97/ele/gens/ and ww

A:Note: published errata appeared in Science 283, 35, 1999

A:Accession: B88102

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-168 <STO>

A:Cross-references: chrchr_11: P00N/A00413.1; PIP:g2915466

C:Genetics.

A:Genes: W09G10.5

A:Map position: 2

A:Accession: A28836
A:Molecule type: DNA
A:Residues: 1-129 <WOO>
A:Cross-references: GR:000940; NID:942921; FTN:CAA5485.1; Hb:J4523
J:Gen Microbiol 132, 3239-3251, 1986
R:Wilde, R.J.; Guest, J.R.
A:Title: Transcript analysis of the citrate synthase and succinate dehydrogenase genes
A:Reference number: 141112; MIM:8800821; PMID:330132
A:Accession: 184546
A>Status: preliminary; translated from GR/EMBL/DBP
A:Molecule type: DNA
A:Residues: 1-129 <RES>
A:Cross-references: GR:MC8940; NID:342942; PTD:AAA2463; FT:455784
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; GC
A:Posn, D.T.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12
A:Reference number: A64720; MIM:9426617; PMID:9278503
A:Accession: H64807
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-129 <RLAT>
A:Cross-references: GR:A060175; GR:000099; NID:9176994; FIDN:AACT3815.1; PII:q1786940
A:Experimental source: Strain K-12, substrain M61659
C:Genetics:
A:Gene: sdhC; cybA
A:Map position: 17 min
C:Complex: one of two hydrophobic anchor components of the succinate dehydrogenase comp
C:Function:
A:Pathway: tricarboxylic acid cycle
C:Superfamily: succinate dehydrogenase 14k hydrophobic protein
C:Keywords: heme; oxidoreductase; transmembrane protein; tricarboxylic acid cycle
E:32-48/Domain: transmembrane *status predicted <TM>
E:73-89/Domain: transmembrane *status predicted <TM>
E:112-128/Domain: transmembrane *status predicted <TM>

Query Match 2 24: Score 7; DR:1; Length 129,
Best Local Similarity 100.08; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LIAGVLY 38
DB 122 LIAGVLY 128

RESULT 8
B90722
succinate dehydrogenase [imported] - *Escherichia coli* (strain 0157:H7, substrain RMD 05
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 *sequence_revision 18-Jul-2001 *txt_change 05-Aug-2001
C:Accession: B90722
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasAqua, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* 0157:H7 and genome
A:Reference number: A96629; MIM:2155231; PMID:11258796
A:Accession: B90722
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-129 <HAV>
A:Cross-references: PTD:BA000007; PTD:BAH44169; PTD:q13360204; GSPR:GN00154
A:Experimental source: strain 0157:H7, substrain RMD 050652
C:Genetics:
A:Gene: ECG746
C:Superfamily: succinate dehydrogenase 14k hydrophobic protein

Query Match 2 24: Score 7; DR:2; Length 129,
Best Local Similarity 100.04; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 LIAGVLY 38
DB 122 LIAGVLY 128

```

RESULT 9
AB5573
succinate dehydrogenase, cytochrome b556 [imported] - Escherichia coli
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16 Feb 2001 #text_char
C:Accession: AB5573
C:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Slattum, G.;
Miller, J.; Grobbeck, T. J.; Davis, N.W.; Lim, A.; Tatum, J. P.;
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli
A:Reference number: AB5480, MULT.1673455, FMDL146551
A:Accession: AB5573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1129 (870)
A:Cross-references: GB:AL51382; FIDN:G005194; FIDN:AG5504
A:Experimental source: strain 057:H7, substrain EDL934
C:Genetics:
A:Gene: sdhc
C:Superfamily: succinate dehydrogenase 14k hydrophobic prot
Query Match
Best Local Similarity 2.2% Score 71 DB 2 Length 129
Matches 7, Conservative 0, Mismatches 0, Indel
OY 32 LLACLV 38
Db 122 LLACLV 128

RESULT 10
AH0590
succinate dehydrogenase cytochrome b 556 chain [imported] -
C:Species: Salmonella enterica subsp. enterica serovar Typh
A:Note: this species has also been called Salmonella typh
C:Date: 09-Nov-2001 #sequence_revision 09 Nov 2001 #text_cha
C:Accession: AH0590
C:Parkhill, J.; Dougan, G.; James, K. D.; Thomson, N. P.; Prie
th, T. J.; Conerton, P. J.; Cronin, A. J.; Davis, P. J.; Davi
S. J.; Moule, S. J.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M
A>Title: Complete genome sequence of a multiple drug resist
A:Reference number: AB0502; PMID:11677608
A:Accession: AH0590
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1129 (870)
A:Cross-references: GB:AL51382; FIDN:G005194.1; PID:G16501
C:Genetics:
A:Gene: SDY0775
C:Superfamily: succinate dehydrogenase 14k hydrophobic prot
Query Match
Best Local Similarity 2.2% Score 71 DB 2 Length 12
Matches 7, Conservative 0, Mismatches 0, Indel
OY 32 LLACLV 38
Db 122 LLACLV 128

RESULT 11
AB2119
succinate dehydrogenase, cytochrome b556 chain WD291 [import
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_cha
C:Accession: AB2119
C:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.;
Richardson, D.; Ermolova, M. D.; Vamathevan, J.; Bass, S.; Gil
L. P.; Mekalanos, J. T.; Venturi, L.C.; Fraser, C.M.
Nature 406, 455-463, 2000
A>Title: Genome sequence of enterohemorrhagic Escherichia coli
A:Reference number: AB5480, MULT.1673455, FMDL146551
A:Accession: AB5573
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1129 (870)
A:Cross-references: GB:AL51382; FIDN:G005194; FIDN:AG5504
A:Experimental source: strain 057:H7, substrain EDL934
C:Genetics:
A:Gene: sdhc
C:Superfamily: succinate dehydrogenase 14k hydrophobic prot
Query Match
Best Local Similarity 2.2% Score 71 DB 2 Length 129
Matches 7, Conservative 0, Mismatches 0, Indel
OY 32 LLACLV 38
Db 122 LLACLV 128

```


A:Accession: S66974
 A:Molecule type: DNA
 A:Cross-references: GR:AE005174, NID:q1251441, PIRN:AA055287
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SCD:VPS21: YPT51
 A:Cross-references: SGD:S0005615; MIPS:Y0P089c
 A:Map position: 15R
 C:Superfamily: ras transforming protein; translation elongation factor Tu homology
 C:Keywords: GTP binding; membrane trafficking; nucleotide binding; P-loop
 F:14-21/Region: nucleotide-binding motif A (P-loop)
 F:120-123/Region: GTP-binding NKXD motif
 F:151-155/Region: GTP-binding SAK/L motif
 F:20/Binding site: GTP (Lys) *status predicted

Query Match 2.2%; Score 7; DB 2; Length 210;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

OY 233 KTAEEON 239
 |||||
 Db 174 KTAEEON 180

RESULT 16
 T35281
 Probable two component response regulator - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #extl_change 31-Jan-2000
 C:Accession: T35281
 R:Seeger, K.; Harris, P.; Bentley, S.P.; Parkhill, J.; Barrett, R.; Rajadream, M.A.
 submitted to the EMBL data library, July 1999
 A:Reference number: Z21574
 A:Accession: T35281
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-227 <SEE>
 A:Cross-references: EMBL:AL096872; PIRN:CA051249.1; GSPR:GN00070; SCD:VPS21: YPT51
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCD:RHS-S05P7.36c
 C:Superfamily: regulatory protein ccmA, response regulator homology

Query Match 2.2%; Score 7; DB 2; Length 227;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

OY 226 RAQIV 242
 |||||
 Db 209 RAQIV 215

RESULT 17
 D85654
 Probable urease accessory protein B [imported] - Escherichia coli (strain 0157:H7, subst
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #extl_change 14-Sep-2001
 C:Accession: D85654
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, R.; Glasner, J.F.; Rose, P.T.; Mayhew
 Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Dimalaria, E.; Fouts, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: genome sequence of enterohemorrhagic Escherichia coli 0157:H7
 A:Reference number: AB5480; MIMD:21074935; PMID:11206551
 A:Accession: D85654
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-232 <STO>
 A:Cross-references: GB:AE005174, NID:q1251441, PIRN:AA055287.1, GSPR:GN00145; MWCP-215
 A:Experimental source: strain 0157:H7, substrain EDL933
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-232 <STO>
 A:Cross-references: GR:AE005174, NID:q1251441, PIRN:AA055287
 A:Experimental source: strain 0157:H7, substrain EDL933
 C:Genetics:
 A:Gene: ureD_2: ureD

Query Match 2.2%; Score 7; DB 2; Length 24;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 267 WVDGSP 273
 |||||
 Db 156 WVDGSP 162

RESULT 18
 AF3451
 IrgB protease [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #extl_change 1
 C:Accession: AF3451
 R:DeLVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Ma
 Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Radu, S.
 Proc. Natl. Acad. Sci. U.S.A. 99, 4444-4448, 2002
 A:Title: The genome sequence of the facultative intracellular
 A:Reference number: AD3252; PMID:11766688
 A:Accession: AF3451
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-235 <SR>
 A:Cross-references: GR:AE008917; PIRN:AA152777.1; PIR:q17980
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME1596
 A:Map position: 1
 C:Superfamily: ynfK protein

Query Match 2.2%; Score 7; DB 2; Length 23;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 32 LAGVIV 38
 |||||
 Db 100 LAGVIV 106

RESULT 19
 E7143
 Hypothetical protein H7282 - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis sp)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #extl_change 1
 C:Accession: E7143
 R:Fraser, C.M.; Norris, S.J.; Welstock, G.M.; White, O.; Se
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidamb
 ley, L.; Weisman, J.; Smith, H.O.; Wenter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the
 A:Reference number: AF1250; MIMD:98442770; PMID:9665876
 A:Accession: E7143
 A:Status: preliminary; nucleic acid sequence not shown; trans
 A:Molecule type: DNA
 A:Residues: 1-239 <COL>
 A:Cross-references: GR:AE001209; GR:AE005520; NID:q3322547;
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0282
 C:Superfamily: syphilis spirochete hypothetical protein pnp

Query Match 2.2%; Score 7; DB 2; Length 24
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 35 GIVAIL 41

A:Experimental source: strain 0157-H7, substrain P1M0 0509452
 C:Genetics:
 A:Gene: PF81321

Query Match
 Best Local Similarity 2.2% Score 7: DB 2: Length 247:
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 267 WVDQSP1.273
 |||||
 DB 171 WVDQSP1.177

RESULT 25

hypothetical protein PA5520 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #ext_change 31-Dec-2000

C:Accession: G82956

R:Stover, C.K.; Pham, X.Q.; Fwinn, A.L.; Mizoguchi, S.D.; Warringer, P.; Hickey, M.J.; R
 admin, S.; Yuan, Y.; Brody, L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lathig, K.; Lim,
 J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: AB2950; MIMD:26437337; PMID:1094404

A:Accession: G82956

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1,247 ~STO

A:Cross-references: GB AE004364; GR AE004051; NID:9495149; P1M0:AM008905.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:
 A:Gene: PA5520

Query Match
 Best Local Similarity 2.2% Score 7: DB 2: Length 247:
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 70 AAVGELS.76
 |||||
 DB 175 AAVGELS.181

RESULT 26

hypothetical protein H10184 - *Haemophilus influenzae* (strain Rd KW20)
 C:Species: *Haemophilus influenzae*
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #ext_change 29-Sep-1999

C:Accession: A64145

R:Feilschmann, R.D.; Adams, M.D.; White, O.; Clayton, P.A.; Kirkness, E.F.; Kurlavag, A
 ; O'Carroll, J.D.; Scott, J.; Shirley, P.; Liu, L.; Glodok, A.; Kelley, J.M.; Weidman, J
 .D.M.; Brandon, R.G.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Osofsky, N.S.M
 Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A:Title: Whole genome random sequencing and assembly of *Haemophilus influenzae* Rd

A:Reference number: A64000; MIMD:65450630; PMID:7542800

A:Accession: A64145

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-275 <ICR>

A:Cross-references: GB:0132704; GR:142023; NID:1157313; P1M0:AM01853.1; P1D:01573140, 1
 C:Note: best homolog was a hypothetical protein from *Escherichia coli*
 C:Superfamily: conserved hypothetical protein YJ1068c

Query Match
 Best Local Similarity 2.2% Score 7: DB 2: Length 275:
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 19 GHGALVL.25
 |||||
 DB 148 GHGALVL.154

RESULT 27
 C81843
 C:Species: *Neisseria meningitidis*
 C:Date: 05-May-2000 #sequence_revision 31-Mar-2000 #ext_change 31-Mar-2000

C:Accession: C81843

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Church,
 ; Holroyd, S.; Jagers, K.; Leather, S.; Mout, S.; Mangall, K
 Nature 404, 502-509, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Nei*
 A:Reference number: AB1775; MIMD:20222564; PMID:10761914

A:Accession: C81843

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-275 <PAR>

A:Cross-references: GR AL162756; GR AL162964; NID:97480991

A:Experimental source: serogroup A, strain 22491

C:Genetics:
 A:Gene: esd; NMA1519

A:Superfamily: conserved hypothetical protein YJ1068c
 C:Keywords: carboxylic ester hydrolase

Query Match
 Best Local Similarity 2.2% Score 7: DB 2: Length 27:
 Matches 7: Conservative 0: Mismatches 0: Indels 0:

OY 19 GHGALVL.25
 |||||
 DB 148 GHGALVL.154

RESULT 28

G81097
 C:Species: *Neisseria meningitidis*
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #ext_change 31-Mar-2000

C:Accession: G81097

R:Petelin, H.; Sanders, N.J.; Heideberg, J.; Jeffries, A.
 Hickey, E.K.; Hart, D.H.; Salzbach, S.L.; White, O.; Feilschmann,
 R.; H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masti
 Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; M
 A:Title: Complete genome sequence of *Neisseria meningitidis*

A:Reference number: AB1000; MIMD:20175755; PMID:10710307

A:Accession: G81097

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-275 <JET>

A:Cross-references: GB:AE002479; GR:AE002098; NID:97226543; 1

A:Experimental source: serogroup B, strain MC58

C:Genetics:
 A:Gene: NMA1305

A:Superfamily: conserved hypothetical protein YJ1068c

Query Match
 Best Local Similarity 2.2% Score 7: DB 2: Length 27:
 Matches 7: Conservative 0: Mismatches 0: Indels 0:

OY 19 GHGALVL.25
 |||||
 DB 148 GHGALVL.154

RESULT 29

A97226
 C:Species: *Clostridium acetobutylicum*
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #ext_change 14-Sep-2001

C:Accession: A97226

R:McClung, J.; Prelog, G.; O'Connell, M.V.; Matraya, K.S.
 ; Daly, M.J.; Honnelt, G.N.; Koonin, E.V.; Smith, D.R.
 J. Bacteriol. 183, 4874-4878, 2001

A:Title: Genome Sequence and Comparative Analysis of the *Coli*

uncharacterized conserved membrane protein CAC2649 [imported]
 C:Species: *Clostridium acetobutylicum*
 C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #ext_change 14-Sep-2001

A:Reference number: A96900; MIMD:21359325; PMID:21359325
 A:Accession: A97226
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-275 <KUR>
 A:Cross-references: GB:AF001437; FIDN:AAK80566.1; FID:0159256773; OSHEW:ON-0149
 A:Experimental source: Clostridium acetobutylicum A10C824
 A:Genetics:
 A:Gene: CAC2649

Query Match: 2.2%; Score 7; DB 2; Length 275;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 22 ALV001.28
 1111111
 92 ALV001.28 98

RESULT 30
 105778

peroxidase (EC 1.11.1.7), seed coat - soybean (fragment)
 C:Species: Glycine max (soybean)
 C:Date: 23 Apr 1999 #Sequence_Revision 23 Apr 1999 #text_change 03 Mar 2000
 C:Accession: T06778

R: Huangpu, J.; Graham, M.C.; Graham, J.S.
 submitted to the EMBL Data Library, November 1995

A:Reference number: 215805
 A:Accession: T06778

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-283 <HUA>
 A:Cross-references: EMBL:J41657; NID:q1125103; FIDN:AAH4129.1; FID:q1125104

A:Experimental source: cultured Williams 28mghly
 A:Genetics:
 A:Gene: SP04.1

C:Function:

A:Superfamily: catalyzes the oxidation of a donor using hydrogen peroxide as electron acceptor

C:Keywords: heme, iron, metalloprotein, oxidoreductase

F:2129/Binding site: heme iron (His) (axial ligands) #status predicted

F:454/Disulfide bonds: #status predicted

F:157-256/Disulfide bonds: #status predicted

Query Match: 2.2%; Score 7; DB 2; Length 283;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 64 N101KA 70
 1111111
 104 N101KA 110

RESULT 31
 H97074
 uncharacterized proteins, homologs of microcin C7 resistance protein MccP (imported)
 C:Species: Clostridium acetobutylicum
 C:Date: 14-Sep-2001 #Sequence_Revision 14 Sep 2001 #text_change 14 Sep 2001
 C:Accession: H97074

R:Rolling, J.; Breton, G.; Omelchenko, M.V.; Markanova, K.S.; Zeng, X.; Gibson, F.; Lee, J.; Daly, M.T.; Bennett, G.N.; Koonin, E.V.; Smith, D.P.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent Producing Bacterium C12

A:Reference number: A96900; MIMD:21359325; PMID:21359325

A:Molecule type: DNA

A:Residues: 1-288 <KUP>

A:Cross-references: GB:AF001437; FIDN:AAK80566.1; FID:0159256773; OSHEW:ON-0149
 A:Experimental source: Clostridium acetobutylicum A10C824
 A:Genetics:
 A:Gene: CAC1419

Query Match: 2.2%; Score 7; DB 2; Length 288;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 295 AEP008G 401
 1111111
 54 AEP008G 60

RESULT 32
 AC2837

5,10-methylerythroidylglyoxylate reductase (imported) - Agrobacterium tumefaciens (strain C:Species: Agrobacterium tumefaciens
 C:Date: 31 Jan 2002 #Sequence_Revision 11 Jan 2002 #text_change 03 Feb 2002
 C:Accession: AC2837

R:Wood, J.W.; Senofal, J.C.; Kraft, P.; Marks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo

erage, G.; Gillet, W.J.; Grant, C.; Guenther, D.; Kutaydin, T.; Levy, R.; Li, M.; McCI

l; Karp, P.; Komarov, P.; Zhang, S.
 Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Huddler, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, P.W.

A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: A96977; PMID:1194193

A:Accession: AC2837

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-296 <KUR>

A:Cross-references: GB:AF008084; FIDN:AAK4113.1; FID:q17740585; OSHEW:ON00186

A:Experimental source: strain C58 (Dupont)

A:Genetics:

A:Gene: metP

A:Maf protein - involved in iron-sulfur cluster assembly

C:Superfamily: 5,10-methylerythroidylglyoxylate reductase (FADH2)

Query Match: 2.2%; Score 7; DB 2; Length 296;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 63 ON101KA 69
 1111111
 211 ON101KA 217

RESULT 33
 JX0209
 Lectin, rabbit skeletal muscle, specific for mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 26 Jun 1992 #Sequence_Revision 30 Jun 1992 #text_change 20 Aug 1999
 C:Accession: JX0209; PX0009

R:Sato, M.; Kawakami, K.; Osawa, T.; Toyoshima, S.
 J. Biochem. 111, 341-346, 1992

A:Title: Molecular cloning and expression of cDNA encoding a galactose/N-acetylglact

A:Reference number: JX0209; MIMD:0236033; PMID:1187794

A:Accession: JX0209

A:Molecule type: mRNA

A:Residues: 1-304 <SAT>
 A:Cross-references: GB:AF066763; FIDN:AAK22573.1; FID:04349461
 R:Yoda, S.; Sato, M.; Toyoshima, S.; Osawa, T.
 J. Biochem. 104, 600-605, 1988
 A:Title: Purification and characterization of a lectin-like molecule specific for gal

A:Reference number: PX0009; MIMD:0197865; PMID:3241002

A:Molecule type: protein

A:Residues: 182-120,137, X',139,171-52A
 C:Superfamily: hepatic lectin C-type lectin homology
 C:Keywords: glycoprotein; lectin; macrophage; transmembrane protein
 F:16-61/Domain: transmembrane #status predicted <TRA>
 F:173-296/Domain: C-type lectin homology <LCH>
 F:174-196/Domain: C-type lectin homology (Asu) (covalent) #status predicted

Query Match: 2.2%; Score 7; DB 2; Length 304;
 Best Local Similarity 100.0%; Pred. No. 58;

Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 236 EONFLO 242
 1111111
 DB 215 EONFLO 221

RESULT 34

H81349

probable periplasmic protein Cj0784 [imported] - Campylobacter jejuni (strain NCTC 11168)
 C:Species: Campylobacter jejuni
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002

C:Accession: H81349
 R:Barthell, J.; Wren, R.W.; Mungall, K.; Kettley, J.M.; Church, P.C.; Rasham, S.; Chittil, C.W.; Quill, M.; Rajandream, M.A.; Fothergill, F.M.; Varvill, A.; Whitehead, S.; Bartel, Nature 403, 665-668, 2000

A:Title: The genome sequence of the food borne pathogen Campylobacter jejuni reveals hyf
 A:Reference number: A81250, MIM:20150912, PMID:10688204

A:Accession: H81349
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-304 <PAR>
 A:Cross-references: GB:AL139076, GB:AL111168, NID:9646126, PIR:969682, PIR:969682

A:Experimental source: serotype O2, strain NCTC 11168
 C:Genetics:
 A:Gene: Cj0784

Query Match
 Best Local Similarity 2.2% Score 7: DB 2: Length 304,
 100.0% Pred. No. 58;

Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 236 EONFLO 242
 1111111
 DB 223 EONFLO 229

RESULT 35

E97614

5,10-methylenetetrahydrofolate reductase (FADH2) (EC 1.7.99.5) [imported] - Agrobacterium
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002

C:Accession: E97614
 R:Frederick, B.; Hinkle, G.; Gitting, S.; Miller, N.; Blanchard, M.; Vuturolo, B.; Goldman, A.; Liu, P.; Wollan, C.; Allington, M.; Toughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359, PMID:11743194

A:Accession: E97614
 A:Status: preliminary
 A:Molecule type: DNA

A:Residues: 1-312 <KOR>
 A:Cross-references: GB:AE007866, PIR:915157268, GSI:000000169

C:Genetics:
 A:Gene: ACK_C_3850

A:Map position: circular chromosome
 C:Superfamily: 5,10-methylenetetrahydrofolate reductase (FADH2)

C:Keywords: oxidoreductase

Query Match
 Best Local Similarity 2.2% Score 7: DB 2: Length 312,
 100.0% Pred. No. 60;

Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 63 ONITOLK 69
 1111111
 DB 227 ONITOLK 233

RESULT 36

F71957

ABC transporter, permease - Helicobacter pylori (strain 199)
 C:Species: Helicobacter pylori
 A:Variety: strain 199

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 11-Jan-2000

C:Accession: F71957

R:Alm, R.A.; Ling, I.S.L.; Mott, D.L.; King, B.L.; Brown, E.L.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Ling, G.; Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolate
 A:Reference number: A71800, MIM:20150912, PMID:9923682

A:Accession: F71957
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-338 <ARN>

A:Cross-references: GB:AE001461, GB:AE001459, NID:94154749, I
 A:Experimental source: strain 199

C:Genetics:
 A:Gene: jhp0236

C:Superfamily: oligopeptide permease protein oppB

Query Match
 Best Local Similarity 2.2% Score 7: DB 2: Length 330,
 100.0% Pred. No. 64;

Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 251 FSWMGLS 257
 1111111
 DB 210 FSWMGLS 216

RESULT 37
 C64551
 oligopeptide ABC transporter, permease protein - Helicobacter
 C:Species: Helicobacter pylori
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Aug-1997

C:Accession: C64551
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Solt, Peterson, S.; Loftus, B.; Richardson, D.; Deason, R.; Khakhria, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujita, Nature 388, 539-547, 1997

A:Authors: Walling, E.; Hayes, W.S.; Fiedorowicz, M.; Karp, P.
 A:Title: The complete genome sequence of the gastric pathogen
 A:Reference number: A64520, MIM:20150912, PMID:9252185

A:Accession: C64551
 A:Status: preliminary; nucleic acid sequence not shown; trans

A:Molecule type: DNA
 A:Residues: 1-338 <TOM>

A:Cross-references: GB:AE000544, GB:AE000511, NID:92433477
 C:Genetics:

A:Start codon: GTG
 C:Superfamily: oligopeptide permease protein oppB

Query Match
 Best Local Similarity 2.2% Score 7: DB 2: Length 330,
 100.0% Pred. No. 64;

Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 251 FSWMGLS 257
 1111111
 DB 210 FSWMGLS 216

RESULT 38
 T05723
 peroxidase (EC 1.11.1.7) precursor, seed coat - soybean
 C:Species: Glycine max (soybean)
 C:Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 09-Apr-1999

C:Accession: T05723
 R:Gijzen, M.
 Plant J. 12, 991-998, 1997

A:Title: A deletion mutation at the cf. locus causes low seed
 A:Reference number: Z15434, MIM:20150912, PMID:9418041

A:Accession: T05723
 A:Status: translated from GR/PM470040

A:Molecule type: DNA
 A:Residues: 1-352 <GID>

A:Cross-references: EMBL:AF014567, NID:4242665, PIR:AA9677
 A:Experimental source: strain Harosoy 61
 C:Genetics:

A:Gene: bp
 A:Introns: 73/3: 137/3: 193/1
 C:Superfamily: peroxidase
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase
 E:1-27/Domain: signal sequence #status predicted <SIG>
 E:28-352/Protein: steel coat peroxidase #status predicted <MA>
 E:47-117/Disulfide bonds: #status predicted
 E:64/Active site: Arg #status predicted
 E:68,195/Binding site: heme iron (His) (axial ligands) #status predicted
 E:70-75/Disulfide bonds: #status predicted
 E:123-325/Disulfide bonds: #status predicted
 E:326-331/Disulfide bonds: #status predicted
 Query Match: 2.2%, Score 7, DB 2, Length 352,
 Best Local Similarity 100.0%, Pred. No. 67,
 Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0,
 64 NTPOLKA 70
 |||||
 DK 170 NTPOLKA 176

RESULT 39
 E:4837
 Probable methyltransferase b0967 Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C:Date: 10 Sep 1999 #sequence_revision 10 Sep 1999 #text_change 03 Mar 2002
 C:Accession: F64837
 R:Blattner, P.P., Plunkett III, G., Plaut, C.A., Petro, N.T., Rife, L.V., Riley, M., et al.
 A:Rice, D.J., Mau, B., Shao, Y.
 Science 277, 1453-1457, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64740; MIM:9742617; PMID:9778502
 A:Accession: F64837
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1367 <BLAT>
 A:Cross-references: GR:AF006158; GR:066066; MIM:q1787189; PIRN:AA074063; PIRN:q1787203
 A:Experimental source: strain K-12, substrain MG1655
 C:Superfamily: Probable methyltransferase b0967

Query Match: 2.2%, Score 7, DB 1, Length 367,
 Best Local Similarity 100.0%, Pred. No. 69,
 Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0,
 23 LVQJLS 29
 |||||
 DB 95 LVQJLS 101

RESULT 40
 A:85624
 Probable oxidoreductase 21319 [imported] Escherichia coli (strain O157:H7, substrain B)
 C:Species: Escherichia coli
 C:Date: 14 Feb 2001 #sequence_revision 14 Feb 2001 #text_change 14 Dec 2001
 C:Accession: AB5624
 R:Berni, N.T., Plunkett III, G., Puriano, V., Mau, B., Glasner, J.P., Rose, D.T., Mayhew
 Miller, L.J., Grobeck, E.T., Davis, N.W., Lam, A., Dimalanta, E., Palamovskis, K., Apostol
 Naure 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MIM:21074935; PMID:11206553
 A:Accession: AB5624
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1367 <STO>
 A:Cross-references: GR:AE005174; MIM:q1251172; PIRN:AA055453.1; SSTRB:280145; SW63.018
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Keywords: enterohemorrhagic
 A:Gene: 21319
 C:Superfamily: Probable methyltransferase b0967
 Query Match: 2.2%, Score 7, DB 2, Length 367,
 Best Local Similarity 100.0%, Pred. No. 69,
 Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0,
 23 LVQJLS 29
 |||||
 DB 95 LVQJLS 101

Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0,
 23 LVQJLS 29
 |||||
 DB 95 LVQJLS 101

RESULT 41
 E:00760
 Probable oxidoreductase b0967 [imported] Escherichia coli (strain O157:H7, substrain B)
 C:Species: Escherichia coli
 C:Date: 14 Feb 2001 #sequence_revision 14 Feb 2001 #text_change 14 Dec 2001
 C:Accession: G90760
 R:Hayashi, T., Makino, K., Chishiki, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.
 Yashima, N., Yashima, T., Kubara, S., Shiba, T., Hattori, M., Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
 A:Reference number: AB0629; MIM:2154231; PMID:11258796
 A:Accession: G90760
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1367 <HAY>
 A:Cross-references: GR:BA000007; PIRN:BA03474.1; PIR:q1369511; SSTRB:GN00154
 A:Experimental source: strain O157:H7, substrain RMD 0509952
 C:Keywords: enterohemorrhagic
 A:Gene: EC01051
 C:Superfamily: Probable methyltransferase b0967

Query Match: 2.2%, Score 7, DB 2, Length 367,
 Best Local Similarity 100.0%, Pred. No. 69,
 Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0,
 23 LVQJLS 29
 |||||
 DB 95 LVQJLS 101

RESULT 42
 D64101
 Influenza A-like virus (H2N2) Haemophilus influenzae (strain PD KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18 Aug 1995 #sequence_revision 18 Aug 1995 #text_change 18 Jun 1999
 C:Accession: D64101
 R:Flersbøll, R., Adams, M.D., Miller, G., Clayton, F.A., Kinkness, E.F., Kurlavage
 J. Gocayne, J.D., Scott, J., Shirley, R., Liu, L., Glodok, A., Kelley, J.M., Weidman
 D.M., Brandon, R., Pile, L.D., Fritchman, J.L., Geoghegan, N.S.M
 Science 269, 496-512, 1995
 A:Authors: Guelm, G., Merviel, L.A., Smith, K.V., Fraser, C.M., Smith, H.G., Venter
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: D64101
 A:Accession: D64101
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1368 <TIG>
 A:Cross-references: GR:U52772; GR:243843; MIM:q1574918; PIRN:AA022700.1; PIR:q177420
 C:Superfamily: glutamate 5-kinase
 C:Keywords: phosphotransferase; protein biosynthesis
 Query Match: 2.2%, Score 7, DB 2, Length 368,
 Best Local Similarity 100.0%, Pred. No. 69,
 Matches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0,
 37 LVAILVQ 43
 |||||
 DB 154 LVAILVQ 160

RESULT 43
 A:95347
 Probable methyltransferase b0967 [imported] Shiga toxin mellei (STC)
 C:Species: Shiga toxin mellei
 C:Date: 24 Aug 2001 #sequence_revision 24 Aug 2001 #text_change 14 Sep 2001
 C:Accession: A95347

```

Query Match      2.4%; Score 7; 38 2; Length 376,
Best Local Similarity 100.0%; Pred. No. 71;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Oy 147 LKAAYGE 143
      | | | | |
Db 209 LKAAYGE 275

```

19677(

Query Match	2.2%	Score 7;	DB 2;	
Best Local Similarity	100.0%	Prod. No. 73;		
Matches	7;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
QY	241	LQLOQTSR	247	
DB	37	LQLOQTSR	43	

D8690C

AHC transporter permease protein ywiH [imported] - *Lactococcus lactis* subsp. *lactis* (strain 7:1species: *Lactococcus lactis* subsp. *lactis*)

Query Match	2.74%	Score 71	DB 2	Length 34
Best Local Similarity	100.00%	Prod. No. 74		
Matches	71	Conservative	0	Maximal Res
QY	256	LSIDNOE	262	
Db	88	LSIDNOE	94	

Job time : 30 secs

Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2002, 11:07:40, Search time 17 seconds

(without alignments)
792,930 Million cell updates/sec

Title: US-09-831-458A-12

Sequence: 1 MSDSKPPVQVQGLLGLGSLH PPTVNVYICKKPAACFIDE 325

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 5

Total number of hits satisfying chosen parameters: 7227

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2.8	956	AD19_HUMAN	Q9H013 Homo sapien
2	2.5	214	FYRE_VIRCH	V9KVD5 Virbio chel
3	2.2	129	DHSC-ECOL1	P10446 Escherichia
4	2.2	129	DHSC-SALTY	O8XG40 Salmonella
5	2.2	179	OPF3_HUMAN	Q9H6K4 Homo sapien
6	2.2	210	YF51_YEAST	P36017 Saccharomyc
7	2.2	239	Y283_TREPA	O83306 Treponema p
8	2.2	275	YAIM_HAEIN	P44356 Haemophilus
9	2.2	304	MMGL_MOUSE	P49300 Mus musculu
10	2.2	367	PROB_PASMU	O9C1J5 Pasteurella
11	2.2	368	PROB_HAEIN	P43763 Haemophilus
12	2.2	390	DXP_FUSNN	O81622 Fusobacteri
13	2.2	396	YCCW-ECOL1	P75676 Escherichia
14	2.2	471	PTMB_BACST	P50852 Bacillus st
15	2.2	497	DHAL-ASPG	P41751 Aspergillus
16	2.2	501	CSD2_DROME	O9VMT6 Drosophila
17	2.2	505	TCMO-POPKI	O43054 Populus kit
18	2.2	527	IMA_VYES	O23478 Lycopersic
19	2.2	548	LACL_PHLRA	O01679 Phlebia rad
20	2.2	583	FRIZ_GROVI	O24760 Drosophila
21	2.2	610	PTMA_BACST	P42956 Bacillus su
22	2.2	708	ABBI_MOUSE	O9QXJ1 Mus musculu
23	2.2	710	ABBI_HUMAN	O00213 Mus sapien
24	2.2	715	FRF2_CANAL	O13354 Candida alb
25	2.2	720	TCME_HUMAN	O43548 Homo sapien
26	2.2	759	ITB3_BOVIN	F32552 Bos taurus
27	2.2	769	ITB3_HUMAN	P05107 Homo sapien
28	2.2	769	ITB3_PIG	F53714 Sus scrofa
29	2.2	778	ITR1_HUMAN	F26010 Homo sapien
30	2.2	804	SCY1_YEAST	P53009 Saccharomyc
31	2.2	804	YBHF-ECOL1	P77504 Escherichia
32	2.2	806	ITB7_MOUSE	P26011 Mus musculu
33	2.2	991	CLK1_DROME	O03445 Drosophila

34	2.2	3414	1	P04G_TREBH
35	1.8	84	1	BECL_HUMAN
36	1.8	104	1	KK21_PORPU
37	1.8	123	1	C59A_MOUSE
38	1.8	124	1	R121_SINY3
39	1.8	125	1	AWA3_TRLAB
40	1.8	129	1	C59B_MOUSE
41	1.8	129	1	CYGD_RHOGE
42	1.8	132	1	KL7A_PIG
43	1.8	140	1	Z049_XENLA
44	1.8	141	1	H0A1_PORCA
45	1.8	141	1	H0F1_DRECA

ALIGNMENTS

RESULT 1
AD19_HUMAN
ID AD19_HUMAN STANDARD: PRT: 956 AA.
AC Q9H013; Q9H02; Q9H2L5;
DT 15-JUN-2002 (Ref. 41, Created)
DT 15-JUN-2002 (Ref. 41, Last sequence update)
DT 15-JUN-2002 (Ref. 41, Last annotation update)
DE ADAM 19 precursor (EC 3.4.24.-) (A disintegrin and meta-
domain 19) (Mellitin beta) (Metalloprotease and disintegr-
DE antigen marker) (MADAM).
GN ADAM19 OR MTNB OR FKSG34.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut-
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Hu-
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RA Wang Y.-G., Gong L.;
RT "Identification of FKSG34, a novel human gene encoding a
RT metalloprotease-disintegrin mellitin beta."
RL Submitted (DEC-2000) to the FMB/Genbank/EMBL data-
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM B).
RA TISSUE=Lymph node;
RC MEDLINE=20346928; Pubmed 10847142;
RA Fritzsche J., Moser M., Faust S., Jekker A., Huettner R.,
RA Krenitz M.;
RT "Molecular cloning and characterization of a human meta-
RT disintegrin a novel marker for dendritic cell different-
RL Hlood 96:732-739(2000)."
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM B).
RA TISSUE=Dendritic cell;
RC MEDLINE=21094672; Pubmed 1152564;
RA Wei P., Zhao Y.-G., Zhang L., Kolon S., Sang Q.-X.A.;
RT "Expression and enzymatic activity of human disintegrin
RT metalloprotease ADAM19/mellitin beta."
PL Biochem. Biophys. Res. Commun. 290:744-755(2001).
RN [4]
RP SEQUENCE OF 100-956 FROM N.A. (ISOFORM A).
RA Xu R., Cai Y., Ying B., Wang F., Xu T., Zhao S., Li C.;
RT "Partial sequence of Homo sapiens ADAM19."
RL Submitted (MAR-1999) to the FMB/Genbank/EMBL data-
RN [5]
RP FUNCTION: PARTICIPATES IN THE PROTEOLYTIC PROCESSING
RN NEUREGULIN ISOFORMS WHICH ARE INVOLVED IN NEUREGENE-
RN SNAPIOGENESIS, SUGGESTING A REGULATORY ROLE IN GLIA
RN CLEAVES ALPHA-2 MACROGLOBULIN. MAY BE INVOLVED IN
RN DIFFERENTIATION AND/OR OSTEOBLAST ACTIVITY IN BONE
RN (SIMILARITY).
CC 1- COFACTOR: BINDS ONE 21N⁺ ION (BY SIMILARITY).
CC 1- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC 1- ALTERNATIVE PRODUCTS: 2 isoforms: A (shown here) and
CC Produced by alternative splicing.
CC 1- TISSUE SPECIFICITY: EXPRESSED IN MANY NORMAL ORGANES
CC SEVERAL CANCER CELL LINES.
CC 1- INDUCTION: BY 1,25(OH)2D3 IN MONOCYTES.

CC	-1	PTM THE PRECURSOR IS CLEAVED BY A FUFIN ENDOPEPTIDASE (BY SIMILARITY).	REFNOMS: TO PEPTIDASE FAMILY M12B.
CC	-1	SIMILARITY. CONTAINS 1 EGF-LIKE DOMAIN	
CC	-1	SIMILARITY. CONTAINS 1 DISINTEGRIN DOMAIN	
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.1sb-slb.ch/announce/ or send an email to license@1sb-slb.ch).	
CC	CC		
DR	EMBL: AF326918:	AAG50282.1	-
DR	EMBL: Y13786:	CAC20585.1	-
DR	EMBL: AF113137:	AAK07852.1	-
DR	EMBL: AF134707:	AAF22162.1	-
DR	HSSP: P18619:	1EVL	
DR	Genew: BCNC:197:	ADAM19	
DR	MIM: 603640:	-	
DR	Interpro: IPR001762:	Disintegrin	
DR	Interpro: IPR000561:	EGF-like	
DR	Interpro: IPR001818:	Matrixin	
DR	Interpro: IPR002870:	Ref-M12b-Propep.	
DR	Interpro: IPR001540:	Ref-M12b-Propep.	
DR	Interpro: IPR000130:	Zn_MTPProtease.	
DR	Pfam: PF00200:	disintegrin_1	
DR	Pfam: PF01421:	Repolysing_1	
DR	Pfam: PF15623:	Pop.M12b-Propep_1	
DR	PRINTS: PR00289:	Disintegrin	
DR	Problem: PD000664:	Disintegrin_1	
DR	SMART: SM00050:	DISIN_1	
DR	PROSITE: PS00215:	ADAM_MPRO_1	
DR	PROSITE: PS00546:	CYSTEINE_SWITCH; FALSE_NEG	
DR	PROSITE: PS00427:	DISINTEGRIN_1; FALSE_NEG.	
DR	PROSITE: PS00214:	DISINTEGRIN_2; 1	
DR	PROSITE: PS00022:	EGF_1; FALSE_NEG.	
DR	PROSITE: PS01186:	EGF_2; 1	
DR	PROSITE: PS00142:	ZINC_PROTEASE; 1	
KW	Hydrolase, Metalloprotease, Zinc, Signal, Glycoprotein, Cysteine, Transmembrane, EGF-like domain, SH3-binding; Alternative splicing.		
ET	SIGNAL	1	25
ET	PROPEP	26	203
ET	CHAIN	204	956
ET	DOMAIN	204	700
ET	TRANSMEM	701	721
ET	DOMAIN	722	956
ET	DOMAIN	204	410
ET	DOMAIN	416	502
ET	DOMAIN	435	438
ET	DOMAIN	503	660
ET	DOMAIN	651	663
ET	SITE	834	840
ET	SITE	839	845
ET	SITE	133	133
ET	SITE	346	346
ET	ACT_SITE	347	347
ET	METAL	350	350
ET	METAL	356	356
ET	DISULFID	321	404
ET	DISULFID	361	388
ET	DISULFID	475	482
ET	DISULFID	655	665
ET	DISULFID	659	671
ET	DISULFID	673	682
ET	CARBHYD	145	145
ET	CARBHYD	445	445
ET	CARBHYD	448	448
ET	CARBHYD	646	646
ET	VARIANT	903	956
ET	KPKAKHSFLVPA -> FPEYRQRAAMTSRT (IN ISOFORM B).		

```

FT CONFLICT 4 4 C->S (IN REF. 2) AND 4).
FT CONFLICT 32 33 SK->R (IN REF. 2 AND 3).
FT CONFLICT 558 558 V->E (IN REF. 2 AND 3).
FT CONFLICT 623 623 N->D (IN REF. 2 AND 3).
SQ SEQUENCE 956 AA; 105038 MW; 8473810FA0418B12 CR64;

Query Match: 7.0%; Score 9; DB 1; Length 956;
Best Local Similarity 100.0%; Pred No 0.79;
Matches 9, Conservative 6, Mismatches 0, Indels 0, Gaps 0.

QY 34 AGYVALIV 42
|||||
Db 706 AGYVALIV 714

```

Query Match	Best Local Similarity	2.5% Score 8	DR 1	Length 214
Matches	8	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0

QY 44 LAGVLVAI 40
 DB 148 LAGVLVAI 155

RESULT 4
 DHSC_ECOLI STANDARD: PRT: 129 AA
 AC P10446;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Succinate dehydrogenase cytochrome b-556 subunit.
 GN SMC OR CYBA OR B0721 OR Z0875 OR EGS0746.
 OS Escherichia coli, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 NX NCBI_TaxID=562, 83374;
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=84307466; PubMed=6943359;
 RA Wood D., Darlison M.G., Wilde R.J., Guest J.R.,
 RT "Nucleotide sequence encoding the flavoprotein and hydrophobic
 RL subunits of the succinate dehydrogenase of Escherichia coli";
 RA Blochem. J. 222:519-534(1984).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=88009821; PubMed=3309132;
 RA Wilde R.J., Guest J.R.;
 RT "Transcript analysis of the citrate synthase and succinate
 RL dehydrogenase genes of Escherichia coli K12";
 RA J Gen Microbiol. 132:3239-3251(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Poth C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Godden M.A., Rose B.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RA Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=97061202; PubMed 8905232;
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuho S., Miki T., Mischuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
 RA Sanei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horiuchi T.;
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
 RL corresponding to the 12.7-28.0 min region on the linkage map";
 RA DNA Res. 3:137-155(1996).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / FDI933 / ATCC 700927;
 RX MEDLINE=21079435; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postel G., Hackett J., Klink S., Boulton A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamotis K.,
 RA Apodaca I., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RA Nature 409:529-533(2001).
 RL [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RIMD 050952;
 RX MEDLINE=21156231; PubMed=11258796;

RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K.
 RA Han C.-G., Ohtsubo E., Nakayama K., Murala T., Tanaka M.
 RA Iida T., Takami H., Honda T., Sasaki G., Ogasawara N.,
 RA Kohara S., Shiba T., Hattori M., Shiragawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia
 RL O157:H7 and genomic comparison with a laboratory strain";
 RA DNA Res. 8:11-22(2001).
 RN [7]
 RP IDENTIFICATION OF PROTEIN.
 RA Murakami H., Kika K., Oya H., Aizaki Y.;
 RT "The Escherichia coli cytochrome b556 gene, cybA, is ass-
 RL in the succinate dehydrogenase gene cluster";
 RA FEMS Microbiol. Lett. 30:307-311(1985).
 RN [8]
 RP MUTAGENESIS OF HISTIDINE RESIDUES.
 RX MEDLINE=98189991; PubMed=9521746;
 RA Vibart C.P., Gerschlager G., Nishimura K., Kita K., Dennis R.;
 RT "Localization of histidine residues responsible for heme
 RL ligation in cytochrome b556 of complex II (succinate dehy-
 CC oxidoreductase) in Escherichia coli";
 CC Biochemistry 37:4148-4159(1998).
 CC [9]
 CC FUNCTION. MONO-HEME CYTOCHROME OF THE SUCCINATE DEHY-
 CC COMPLEX.
 CC [10]
 CC PATHWAY. Tetracycline and cycle.
 CC [11]
 CC SUBUNIT: PART OF AN ENZYME COMPLEX CONTAINING FOUR S-
 CC FLAVO-PROTEIN, AN IRON-SULFUR, CYTOCHROME B-556, AND AN
 CC ANCHOR PROTEIN.
 CC [12]
 CC SIMILARITY: BELONGS TO THE CYTOCHROME B556 FAMILY.
 CC [13]
 CC This SWISS-PROT entry is copyright. It is produced through
 CC between the Swiss Institute of Bioinformatics and the
 CC the European Bioinformatics Institute. There are no re-
 CC use by non-profit institutions as long as its content
 CC modified and this statement is not removed. Usage by
 CC entities requires a license agreement (See [http://www.bbi-
 CC](http://www.bbi-

 CC or send an email to license@bbs.sib.ch)

Query Match: 2.2%; Score 7; DB 1; Length 12
 Best Local Similarity 100.0%; Fred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0

QY 33 LAGVLV 38
 DB 123 LAGVLV 128

RESULT 4
 DHSC_SALTY STANDARD: PRT: 129 AA.
 ID DHSC_SALTY

OY 223 QVRAOL 229
 DB 157 QVRAOL 163

RESULT 6

YPS1_YEAST STANDARD: PRT: 210 AA.
 AC P46017;
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE GTP-binding protein YPS1/YPS21.
 GN YPS1 OR YPS21 OR YOK89C OR YOK3154C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetalia; Saccharomycetaceae; Saccharomycetaceae.
 NCBI_TaxID:4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE:94216368; PubMed:8163546,
 RA Singer-Krueger H., Stenmark H., Duesterhoeft A., Philippson P.,
 RA Yoo J.-S., Gallwitz D., Zerial M.;
 RT "Role of three rab5-like GTPases, Yps1p, Yps2p, and Yps3p, in the
 RT endocytic and vacuolar protein sorting pathways of yeast.";
 RL J. Cell Biol. 125:283-298(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE:94185635; PubMed:8137814;
 RA Horadzowsky B.F., Busch G.R., Emr S.D.;
 RT "YPS21 encodes a rab5-like GTP binding protein that is required for
 RT the sorting of yeast vacuolar proteins.";
 RL EMBO J. 13:1297-1309(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE:97344368; PubMed:9200815;
 RA Voss H., Benes V., Andrade M.A., Valencia A., Bachmann S., Teodori C.,
 RA Schwager C., Paces V., Sander C., Ansorge W.;
 RT "DNA sequencing and analysis of 130 kb from yeast chromosome XV.";
 RL Yeast 13:655-672(1997).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 1-174.
 RA MEDLINE:20221439; PubMed:13756108,
 RA Esters H., Alexandrov K., Constantinescu A.-T., Grody P.S.,
 RA Schofield A.J.;
 RT "High-resolution crystal structure of S. cerevisiae Yps1(DeltaC15)-
 RT GppNHp, a small GTP-binding protein involved in regulation of
 RT endocytosis.";
 RL J. Mol. Biol. 298:111-121(2000).
 CC -1- FUNCTION: REQUIRED FOR TRANSPORT IN THE ENDOCYTIC PATHWAY AND
 CC FOR CORRECT SORTING OF THE VACUOLAR HYDROLASES SUGGESTING A
 CC POSSIBLE INTERSECTION OF THE ENDOCYTIC WITH THE VACUOLAR SORTING
 CC PATHWAY. ACT EITHER AT THE STAGE OF DELIVERY OF PROTEINS FROM THE
 CC GOLGI TO THE LATE ENDOSOME OR FROM THE LATE ENDOSOME TO THE
 CC VACUOLE.
 CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC RAB5 HOMOLOGUE.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by a commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X76173; CAA53769.1;
 CC EMBL: Z29338; CAA82543.1;
 CC EMBL: X94335; CAA64010.1;
 CC EMBL: Z74997; CAA99285.1;
 CC PIR: S40001; S40001.
 CC PIR: S43399; S43399.
 CC PDB: 1EK0; 17-APR-00.

DR Sbj: S005615; YPS21.
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR001806; Ras_Trimstim.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00071; Ras_1.
 DR PRINTS: PR00449; RASINSTRFMMNT.
 DR SMART: SM00175; RAB_1.
 DR TIGRFAMs: TIGR00231; small_gtp.
 KW Protein transport; GTP-binding; Lipoprotein; Prenylation;
 KW 3D-structure.
 FT NP_BIND 14 21 GTP.
 FT NP_BIND 62 66 GTP.
 FT NP_BIND 120 123 GTP.
 FT DOMAIN 36 44 EFFECTOR REGION (POTENTIAL).
 FT LIPID 208 208 GERANYL-GERANYL (BY SIMILAR).
 FT LIPID 210 210 GERANYL-GERANYL (BY SIMILAR).
 FT MUTAGEN 21 21 SEVERE VACUOLAR PROTEIN DEFECT.
 FT
 SO SEQUENCE 210 AA; 23081 MW; 4471A0090C3F09 C6764;

Query Match

Best Local Similarity 100.0%; Ident. No. 20;
 Matches 7; Conservative 6; Mismatches 0; Indels

OY 233 KTAEEON 239
 DB 174 KTAEEON 180

RESULT 7

Y282_TREPA STANDARD: PRT: 219 AA.
 AC 083306;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein TP0282.
 GN TP0282.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 NCBI_TaxID:160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=Nichols;
 RA MEDLINE:98332770; PubMed:945876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sult
 RA Fraser R., Wilson M., Hickey R., Anderson E., Ketchum K.A.,
 RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L.,
 RA Khajak H., Richardson D., Howell J.K., Chidambaram M.,
 RA McDonald L., Artach P., Bowman C., Nelson M.D., Fujita C.,
 RA Hatch R., Horst K., Roberts K., Sankovsky M., Weidman G.,
 RA Venter J.C.;
 RT "Complete genome sequence of Treponema pallidum, the syph
 RT spirochete.";
 RL Science 281:1375-1388(1998).
 CC -1- SIMILARITY: CONTAINS 2 TPR REPEATS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by a commercial
 CC entities requires a license agreement (See [http://www.isb.ch/](http://www.isb.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE001209; AAC65279.1;
 CC DR TIGR: TP0282.
 DR InterPro: IPR001440; TPR.
 KW Hypothetical protein; Repeat; TPR repeat; Transmembrane;
 KW Complete proteome.
 FT TRANSMEM 21 43 Potential.
 FT REPEAT 112 145 TIR 1.
 FT REPEAT 149 182 TIR 2.

SEQUENCE 239 AA, 26415 MW, F483P78934RPF73 CPG64,
 Query Match 2.2%, Score 7, DB 1, Length 239;
 Best Local Similarity 100.0%, Prev. No. 23;
 Patches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

UY 35 GVLVAIL 41
 1111111
 DB 24 GVLVAIL 30

RESULT 8
 YAIM_HAEIN
 ID YAIM_HAEIN STANDARD: PRT, 275 AA
 AC P44556
 DT 01-NOV-1995 (rel. 32, Created)
 DT 01-NOV-1995 (rel. 32, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Protein H11084.
 GN H10184
 OS Haemophilus influenzae.
 OC Bacteria, Proteobacteria, gamma subdivision, Pasteurellales
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KM-20 / ATCC 51907;
 RX MEDLINE=95550630; PubMed=7542800;
 PA Fleischmann P. D., Adams M. D., White O., Clayton R. A., Kirkness P. F.,
 Kerlavage A. R., Bult C. J., Tomb J.-F., Dougherty R. A., Merrick T. M.,
 McKenney K., Sutton G., Fitzhugh W., Fields C. A., Gocayne J. D.,
 Scott J. D., Shirley R., Liu L.-T., Glodek A., Kelley J. M.,
 Weidman J. F., Phillips C. A., Spriggs T., Hedblom F., Cotton M. D.,
 Utterback T. R., Hanna M. C., Nguyen D. T., Saudek J. M., Brandon R. C.,
 Fine L. D., Fritchman J. L., Fritchman N. S. M.,
 Gnehm C. L., McDonald L. A., Small K. V., Fraser C. M., Smith H. O.,
 Venter J. C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RT Science 269:496-512(1995).
 RL [2]
 RN IDENTIFICATION BY MASS SPECTROMETRY.
 RP MEDLINE=20137488; PubMed=10675023;
 RA Langen H., Takacs R., Evans S., Pfeiffer P., Lahn B. W., Wilm F.,
 Gray C., Fountoulakis M.;
 PT "Two-dimensional map of the proteome of Haemophilus influenzae";
 RL Electrophoresis 21:431-439(2000)
 CC -1- SIMILARITY: STRONG, TO E. COLI YAIM AND YERS;
 CC -1- SIMILARITY: STRONG, TO HUMAN ESTERASE P.
 CC -----
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----
 CC EMBL: 032703; AAC13853.1;
 DB TIGR: H10184.
 DR InterPro: IPR000803; EST-Inter-Link;
 DR InterPro: IPR000379; Ser-esterase; 1;
 DR Pfam: PF00705; Esterase; 1;
 KW Complete proteome;
 SO SEQUENCE 275 AA, 31317 MW, /B1P5C041154bD3D CPG64;

Query Match 2.2%, Score 7, DB 1, Length 275;
 Best Local Similarity 100.0%, Prev. No. 26;
 Patches 7, Conservative 0, Mismatches 0, Indels 0, Gaps 0

UY 19 GHGALVI 25
 1111111
 DB 148 GHGALVI 154

RESULT 9
 MMGL_MOUSE
 ID MMGL_MOUSE STANDARD: PRT, 304 AA
 AC P49300;
 DT 01-FEB-1996 (rel. 33, Created)
 DT 01-FEB-1996 (rel. 33, Last sequence update)
 DT 01-NOV-1997 (rel. 35, Last annotation update)
 DE Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (Macrophage
 DE galactose/N-acetylgalactosamine specific lectin) (MMGL).
 GN MGL.
 OS Mus musculus (Mouse).
 OC Eukaryota, Mammalia, Chordata, Vertebrata, Eutelestomi;
 OC Mammalia, Eutheria, Rodentia, Scuriognathia, Muridae, Murinae, Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HEJ;
 RX MEDLINE=92258932; PubMed=1587794;
 PA Sato M., Kawakami K., Osawa T., Toyoshima S.;
 RT "Molecular cloning and expression of cDNA encoding a galactose/N-
 RT acetylgalactosamine-specific lectin on mouse tumoricidal
 RT macrophages.";
 RL J. Biochem. 111:331-336(1992).
 RN [2]
 RC STRAIN=C3H/HEJ; OF 102-179 AND 137-151.
 RX MEDLINE=9379795; PubMed=3441092;
 PA Oda S., Sato M., Toyoshima S., Osawa T.;
 RT "Purification and characterization of a lectin-like molecule specific
 RT for galactose/N-acetyl-galactosamine from tumoricidal macrophages.";
 RL J. Biochem. 104:600-605(1988).
 CC -1- FUNCTION: KETOXINIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE
 CC UNITS MAY PARTICIPATE IN THE INTERACTION BETWEEN TUMORICIDAL
 CC MACROPHAGES AND TUMOR CELLS.
 CC -1- SUBUNIT: HOMO-OLIGOMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II membrane protein.
 CC -1- TISSUE SPECIFICITY: IS EXPRESSED ON THE SURFACE OF ACTIVATED
 CC MACROPHAGES.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -----
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----
 CC EMBL: S16675; AB023171.1;
 DB HSSP: P06734; IHL1.
 DB MGI: M139675; MGI
 DR InterPro: IPR001004; Lectin_C.
 DR Pfam: PF00059; Lectin_C; 1;
 DR SMART: SM0034; CLECT; 1;
 DR PROSITE: PS00615; C-TYPE_LLECTIN_1;
 DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1;
 KW Lectin, Glycoprotein, Transmembrane, Calcium, Signal anchor,
 KW LECTIN, CYTOPLASMIC (POTENTIAL)
 FT DOMAIN 1 45
 FT TRANSMEM 36 56 SIGNAL ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL);
 FT DOMAIN 57 304 EXTRACELLULAR (POTENTIAL);
 FT DOMAIN 172 298 C-TYPE LECTIN (UNOS FORM);
 FT DISULFID 173 184 BY SIMILARITY;
 FT DISULFID 201 296 BY SIMILARITY;
 FT DISULFID 274 286 BY SIMILARITY;
 FT CAPSULE 74 74 N-LINKED (GLYCAN...) (POTENTIAL);
 FT CAPSULE 135 135 N-LINKED (GLYCAN...) (POTENTIAL);
 FT CAPSULE 135 135 N-LINKED (GLYCAN...) (POTENTIAL);
 SO SEQUENCE 304 AA, 34596 MW, 3679CD12C344P5BCC CPG64;

Query Match 2.2%, Score 7, DB 1, Length 304;
 Best Local Similarity 100.0%, Prev. No. 28;

Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 246 EQUNFLQ 242
DB 215 EQUNFLQ 221

RESULT 10

PROB_PASMO
ID PROB_PASMO STANDARD: PRT: 367 AA.
AC Q9CJ05;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Glutamate 5-kinase (EC 2.7.2.11) (gamma-glutamyl kinase) (CK).
GN PROB OR PM1896.
OS Pasteurella multocida.
OC Bacteria: Proteobacteria: gamma subdivision: Pasteurellales: Pasteurella.
OX NCBI_TaxId=747;
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN=pm70;
RA MEDLINE=21145866; PubMed=11248100;
RA May B.-J., Zhang Q., Li L.-I., Paustian M.L., Whitam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida pm70.";
RL Proc Natl Acad Sci U S A 98:3460-3465(2001).
CC -1- FUNCTION: Catalyzes the transfer of a phosphate group to glutamate to form glutamate 5-phosphate which rapidly cyclizes to 5-oxoproline.
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate = ADP + L-glutamate 5-phosphate.
CC -1- PATHWAY: Proline biosynthesis: first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE GLUTAMATE 5-KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announc/> or send an email to license@isb-sib.ch).
CC
CC EMBL: AE006227; AAK03980.1; ALT_INIT.
CC InterPro: IPR001048; Aa_kinase.
CC InterPro: IPR002478; Glu_5kinase.
CC InterPro: IPR002478; PUA.
CC Pfam: PF00696; aakkinase; 1.
CC Pfam: PF01472; PUA; 1.
CC PRINTS: PR00474; GLU5KINASE.
CC SMART: SM00359; PUA; 1.
CC TIGRFAMs: TIGR01027; PROB; 1.
CC PROSITE: PS00903; GLUTAMATE_5_KINASH; 1.
CC TRANSFERASE, KINASE, Proline biosynthesis: Complete proteome.
CC SOURCE: 367 AA; 39894 MW; E62DA632E78FE5 CRC64;

Query Match 2.2%: Score 7; DB 1, Length 367,
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 37 LVAILVQ 43
DB 155 LVAILVQ 161

RESULT 11

PROB_HAETN
ID PROB_HAETN STANDARD: PRT: 368 AA.
AC P43763;
DT 01-NOV-1995 (Rel. 42, Created)
DT 01-NOV-1995 (Rel. 42, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glutamate 5-kinase (EC 2.7.2.11) (gamma-glutamyl kinase)
GN PROB OR H10900.
OS Haemophilus influenzae.
OC Bacteria: Proteobacteria: gamma subdivision: Pasteurella
OC Haemophilus
OX NCBI_TaxId=727;
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN=Fd / KW20 / ATCC 51907;
RA MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kie
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., M
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Goeyn
RA Scott J.D., Shirley K., Liu T.-I., Glodek A., Kelley J.M
RA Widmayer J.F., Phillips C.A., Spriggs T., Heblum E., Cer
RA Osterback T.R., Hanna M.C., Nguyen B.T., Sander D.M., R
RA Fine L.D., Fritchman J.L., Fuhrman J.L., Goodagen N.S.,
RA Guehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smi
RA Venter J.C.;
RT "Whole genome random sequencing and assembly of Haemoph
RL Science 269:496-512(1995).
CC -1- FUNCTION: Catalyzes the transfer of a phosphate group
CC to form glutamate 5-phosphate which rapidly cyclizes
CC to oxoproline.
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate = ADP + L-glut
CC phosphate.
CC -1- PATHWAY: Proline biosynthesis: first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE GLUTAMATE 5-KINASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announc/> or send an email to license@isb-sib.ch).
CC
CC EMBL: U32772; AAC22560.1;
CC TIGR: H10900;
CC InterPro: IPR001048; Aa_kinase.
CC InterPro: IPR001057; Glu_5kinase.
CC InterPro: IPR002478; PUA.
CC Pfam: PF00696; aakkinase; 1.
CC Pfam: PF01472; PUA; 1.
CC PRINTS: PR00474; GLU5KINASE.
CC SMART: SM00359; PUA; 1.
CC TIGRFAMs: TIGR01027; PROB; 1.
CC PROSITE: PS00903; GLUTAMATE_5_KINASH; 1.
CC TRANSFERASE, KINASE, Proline biosynthesis: Complete prote
CC SOURCE: 368 AA; 40134 MW; E32H684A23709831 CRC64;

Query Match 2.2%: Score 7; DB 1, Length 367,
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 37 LVAILVQ 43
DB 154 LVAILVQ 160

RESULT 12

PROB_FUSNN
ID PROB_FUSNN STANDARD: PRT: 490 AA.
AC Q8R622;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-deoxy-D-xylulose 5-phosphate reductoisomerase (EC 1.1.1.26)
DE reductoisomerase (1-deoxyxylulose 5-phosphate reductio
GN DXR OR FN1324.
OS Fusobacterium nucleatum (subsp. nucleatum).

```

CC Bacteria; Fusobacteriales; Fusobacterium
CC NCBI_TaxID=76866;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN ATCC 25586;
CC MEDLINE=21886444; PubMed=11884109;
CC Kapatal V., Anderson I., Teodoro N., Fenchel G., De S. T., Fitch A.,
CC Bhattacharya A., Bartman A., Gardner W., Gueckin G., Zhu L.,
CC Vastava O., Chu L., Kogan Y., Chada O., Goldsman E., Benai A.,
CC Larsen N., D'Souza M., Malinas T., Pusch G., Haselkorn R.,
CC Forstner M., Kyprides N., Overbeck R.;
CC "Genomic sequence and analysis of the oral bacterium Fusobacterium
CC nucleatum strain ATCC 25586."
CC J. Bacteriol. 184:2005-2014(2002)
CC -1- FUNCTION: Catalyzes the NADP-dependent rearrangement and reduction
CC of 1-deoxy-D-xylulose-5-phosphate (DXR) to 2 C-methyl D-erythritol
CC 4-phosphate (MEP) (by similarity).
CC -1- CATALYTIC ACTIVITY: 2 C-methyl-D-erythritol 4-phosphate + NADP(+)
CC = 1-deoxy-D-xylulose 5-phosphate + NADPH.
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; second
CC step.
CC -1- SIMILARITY: BELONGS TO THE DXR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AE010637; AAL95520.1;
CC 1SEQUENCE BIOSYNTHESIS: oxidoreductase; NADP; Complete proteome.
CC NP_BIND 10 NADPH (POTENTIAL).
CC SEQUENCE 390 AA; 43984 MW; 1C3771487885D3ED C6C64;
CC -----
Query Match: 2.2%, Score 7, DB 1, Length 390;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
ID 56 SHODAIY 62
CC [1]
CC 61 SHODAIY 67
CC -----
RESULT 13
ID YCCW_ECOLI STANDARD; PRT: 396 AA.
AC P75876; Q98702;
DT 15-JUN-1998 (Rel. 35, created)
DT 16-OCT-2001 (Rel. 40, last sequence update)
DT 16-OCT-2001 (Rel. 40, last annotation update)
DE Hypothetical protein yccw.
GN YCCW OF P09667
OS Escherichia coli.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Escherichia
CC NCBI_TaxID=562;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN-K12 / MG1655;
CC MEDLINE=97426617; PubMed=9778503;
CC Rhatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Putland V.,
CC Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
CC Mau B., Shao Y.;
CC "The complete genome sequence of Escherichia coli K-12."
CC Science 277:1453-1474(1997).
CC [2]
CC SPOURCE FROM N.A.
CC STRAIN-K12;
CC MEDLINE 97061202; PubMed=8905232;
CC Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,

```

```

CC Kometani K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimura K.,
CC Kimura G., Kitayama M., Maeno K., Masata S., Miki T., Mizobuchi K.,
CC Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
CC Sampei G., Seki Y., Takami H., Takemoto K., Wada C., Yamamoto Y.,
CC Yano M., Horinouchi T.;
CC "A 7.8-kb DNA sequence of the Escherichia coli K-12 genome
CC corresponding to the 3.7-3.9 Mb region on the linkage map."
CC J. Bacteriol. 178:1137-1146(1996).
CC -1- SIMILARITY: BELONGS TO THE DEPO064 FAMILY
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AB000196; AAC74053.1; ALT_INIT.
CC EMBL: J00734; AAA55732.1;
CC ProGene: E013725; YCCW.
CC DR InterPro: IPR002478; PUA.
CC DR InterPro: IPR001250; Prenyl_site.
CC DR InterPro: IPR000051; SAM_bind.
CC DR SMART: SM00359; PUA; 1.
CC KW Hypothetical protein; Complete proteome.
CC SEQUENCE 496 AA; 44454 MW; 0905FCEB627343094 C6C64;
CC -----
Query Match: 2.2%, Score 7, DB 1, Length 396;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC -----
ID 23 LVGLILE 29
CC [1]
CC 124 LVGLILE 130
CC -----
RESULT 14
ID PTMB_BACST STANDARD; PRT: 471 AA.
AC P50852;
DT 01-OCT-1996 (Rel. 34, created)
DT 01-OCT-1996 (Rel. 34, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Pts system, mannitol-specific TIBC component (EIIIC-Mtl) (Mannitol-
DE peucetase TIBC component) (Phosphotransferase enzyme II, BC component)
DE (EC 2.7.1.69) (EII-Mtl).
GN MTLA.
OS Bacillus stearothermophilus.
CC Bacteria; Firmicutes; Bacillales; Geobacillus.
CC NCBI_TaxID=1422;
CC [1]
CC SEQUENCE FROM N.A.
CC STRAIN ATCC 7954;
CC MEDLINE=96421984; PubMed=8824601;
CC Henstra S.A., Tolner R., Howe Purkians P.H., Konings W.N.,
CC Rodilland G.T.;
CC "Cloning, expression, and isolation of the mannitol transport protein
CC from the thermophilic bacterium Bacillus stearothermophilus."
CC J. Bacteriol. 178:5586-5591(1996).
CC -1- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOTRANSFERASE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS). A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE PTS PMAINS CONTAIN THE SUGAR BINDING SITE
CC AND THE TRANSMEMBRANE CHANNEL. THE 11A DOMAIN CONTAINS THE PRIMARY
CC PHOSPHORYL GROUP TO THE 11B DOMAIN WHICH FINALLY TRANSFERS IT TO
CC THE SUGAR.
CC -1- CATALYTIC ACTIVITY: protein N-phosphotransferase + sugar -> protein
CC N-phosphotransferase + sugar phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: CONTAINS 1 PTS EIIA DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PTS EIIB DOMAIN.
CC -----

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U18943; AAC44463.1; -
 DR InterPro: IPR004718; PTS1LC_m1A
 DR Tricfams: TRICF00851; m1A: 1.
 KW Phosphotransferase system: Sugar transport: Transferase:
 KM Transmembrane; Phosphorylation.
 FT DOMAIN 1 339
 FT DOMAIN 1 339
 FT TRANSMEM 30 50
 FT TRANSMEM 54 74
 FT TRANSMEM 92 112
 FT TRANSMEM 143 163
 FT TRANSMEM 219 239
 FT TRANSMEM 273 293
 FT TRANSMEM 294 314
 FT TRANSMEM 319 339
 FT MOD_RES 262 262
 FT MOD_RES 389 389
 SO SEQUENCE 471 AA; 50153 MW; 3BF2F93C8ED968CH CRC64;

Query Match 2.28; Score 7; DB 1; Length 471;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 33 LAGVLA 39
 DB 422 LAGVLA 328

RESULT 15
 ID DHAL_ASPNG STANDARD; PRT: 497 AA.
 AC P41751;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Aldehyde dehydrogenase (EC 1.2.1.3) (AlDH).
 GN ALDA.
 OS Aspergillus niger.
 CC Eukaryota, Fungi, Ascomycota, Pezizomycotina, Eurotiomycetes;
 CC Eurotiiales; Trichocomaceae; trichocomaceae; Aspergillus.
 CC NCBI_TaxID=5061;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90108706; PubMed=2606357;
 RA O'Connell M.J., Kelly J.M.;
 RT *Physical characterization of the aldehyde-dehydrogenase-encoding
 RT gene of Aspergillus niger.*;
 RL Gene 84:173-180(1985)
 CC -1- CATALYTIC ACTIVITY: An aldehyde + NAD(+) + H(2)O = an acid + NADH.
 CC -1- PATHWAY: Ethanol utilization; second step.
 CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M3251; AAA87596.1; -
 DR HSSP: P05091; ICW3.
 DR InterPro: IPR002086; Aldehyde_dehydr.
 DR Pfam: PF00171; aldehyd. 1.
 DR PROSITE: PS00070; ALDEHYDE_DEHYDR_CYS: 1.

DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU: 1.
 KW Oxidoreductase; NAD.
 FT NR_BIND 242 247
 FT ACT_SITE 265 265
 FT ACT_SITE 294 294
 FT ACT_SITE 394 394
 SO SEQUENCE 497 AA; 53809 MW; 4DVA454EDFA50E8 CRC64;

Query Match 2.28; Score 7; DB 1; Length 497;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

UY 230 VIKTAE 236
 DB 186 VIKTAE 192

RESULT 16
 ID CSU2_DROME STANDARD; PRT: 501 AA.
 AC U9VM16;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable cytochrome P450 2b4 (EC 1.14.14.1) (CYP2B4).
 CN CYP2B4 OR CG6081.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; B.
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 CC NCBI_TaxID 7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 KX MEDLINE=20194006; PubMed=10741192;
 RA Adams M.D., Colniker S.E., Holt R.A., Evans C.A., Gonsky
 RA Amandides P.C., Scherer S.E., Li P.W., Hoskins R.A., G.
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Heit
 RA Sutton G.G., Wortman J.P., Yandell M.D., Zhang Q., Chen
 RA Randon R.C., Rogers Y.-H.C., Ritzel P.G., Champe M., J.
 RA Wan K.H., Doyle C., Baxter F.C., Helt G., Nelson G.R., J.
 RA Abell J.F., Abmayr A., An H.-J., Andrews-Plannkuch C.
 RA Balow R.M., Basu A., Baxendale J., Bayraktarolu L., B.
 RA Beeson K.V., Benos P.V., Herman B.P., Bhandari D., Bolis
 RA Berkova P., Berchan M.F., Bock J., Buckstein P., Brod
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A.
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davi
 RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew J.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Durov
 RA Durbin K.J., Evangelista C.C., Foray C., Ferreira S.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., G.
 RA Gluck A., Gong F., Gottlieb J.H., Gu Z., Guan P., Hart
 RA Harris N.L., Harvey D., Heiman J.J., Hernandez J.R., H.
 RA Hostin D., Houston K.A., Howland J., Wei M.-H., Ilegu
 RA Jaitai M., Kalush F., Karpen A., Ko Z., Kennison J.A.
 RA Kimmel P.E., Kodira G.D., Krall C., Kravitz S., Kulp D.
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y.
 RA Liu X., Mattei B., McIntosh T.C., Melrod M.P., Mephorso
 RA Merkulov G., Milshina N.V., Moberg C., Morris J., Mos
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murty D.M., R.
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., P.
 RA Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V.
 RA Reiter K., Remington K., Saunders R.D., Scheeler P.
 RA Shie R.C., Siden-Kiamos I., Simpson M., Skupski M.P.,
 RA Spleer E., Spradling A.C., Stapleton M., Strong R., Su
 RA Svirskis R., Tector C., Turner R., Venter E., Wang A.
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S.
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao
 RA Zheng X.H., Zhong F.N., Zhong W., Zhu X., Zhu S., Zhu
 RA Zibers R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT *The genome sequence of Drosophila melanogaster.*;
 RL Science 287:2185-2195(2000).

```

4P CONCEPTUAL TRANSLATION.
4A Nelson B.:
4L Unpublished observations (SEP-2000).
4C -1- FUNCTION: MAY BE INVOLVED IN THE METABOLISM OF INSECT HORMONES AND
4C IN THE BREAKDOWN OF SYNTHETIC INSECTICIDES (BY SIMILARITY).
4C -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = RH +
4C oxidized flavoprotein + H(2)O.
4C -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
4C (potential).
4C -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY
4C -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous
4C gene model prediction.
4C This SWISS-PROT entry is copyright. It is produced through a collaboration
4C between the Swiss Institute of Bioinformatics and the EMBL outstation -
4C the European Bioinformatics Institute. There are no restrictions on its
4C use by non-profit institutions as long as its content is in no way
4C modified and this statement is not removed. Usage by and for commercial
4C entities requires a license agreement (See http://www.isb-sib.ch/announce/
4C or send an email to license@isb-sib.ch).
4C -----
4C EMBL: AE003609; AAF52235.1; ALT_SEQ.
4C FlyBase: FBgn001668; cyp28a2
4C InterPro: IPR01128; Cytochrome_P450.
4C Pfam: PF00067; P450_1.
4C PRINTS: PR00385; P450.
4C PROSITE: PS00086; CYTOCHROME_P450_1.
4C Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
4C Endoplasmic reticulum; Hypothetical protein.
4C BINDING: 446 446 HEME (BY SIMILARITY).
4C SEQUENCE: 501 AA: 5822 MW: 9486734930AACE CPYC4.
4C -----
4C Query Match: 2.28; Score 7; DB 1; Length 501;
4C Best Local Similarity: 100.0%; Pred. No. 44;
4C Matches: 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
4C -----
4C 88 LITOKAA 94
4C |||||
4C 453 LITOKAA 459
4C -----
4C RESULT 17
4C TM_O_PORKE
4C AC Q43054; STANDARD; PRT; 505 AA.
4C 15-DEC-1998 (Rel. 37, Created)
4C 15-DEC-1998 (Rel. 37, Last sequence update)
4C 15-JUN-2002 (Rel. 41, Last annotation update)
4C Trans-cinnamate 4-monooxygenase (EC 1.14.13.11) (Cinnamic acid
4C 4-hydroxylase) (C4H) (P450C4H) (Cytochrome P450 73).
4C CYP73A16 OR CYP73.
4C OS Populus kirkamienensis (Aspen).
4C Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
4C Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
4C eustosids; I: Malpighiales; Salicaceae; Populus.
4C NCBI_TaxID=3698;
4C [1]
4C SEQUENCE FROM N.A.
4C KX MEDLINE:97141312; PubMed:8987656;
4C Kawai S., Mori A., Shiohara T., Kajita S., Katayama Y., Moroshita N.;
4C "Isolation and analysis of cinnamic acid 4-hydroxylase homologous
4C genes from a hybrid aspen, Populus kirkamienensis";
4C Biosci. Biotechnol. Biochem. 60:1586-1597(1996).
4C -1- FUNCTION: CONTROLS CARBON FLUX TO PIGMENTS ESSENTIAL FOR
4C POLLINATION OR UV PROTECTION, TO NUMEROUS PHYTOALEXINS SYNTHESIS IN
4C BY PLANTS WHEN CHALLENGED BY PATHOGENS, AND TO LIGNINS.
4C -1- CATALYTIC ACTIVITY: Trans-cinnamate + NADPH + O(2) = 4-
4C hydroxycinnamate + NADP(+) + H(2)O.
4C -1- PATHWAY: Phenylpropanoid metabolism; second step.
4C -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
4C -----
4C This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

4C between the Swiss Institute of Bioinformatics and the EMBL outstation -
4C the European Bioinformatics Institute. There are no restrictions on its
4C use by non-profit institutions as long as its content is in no way
4C modified and this statement is not removed. Usage by and for commercial
4C entities requires a license agreement (See http://www.isb-sib.ch/announce/
4C or send an email to license@isb-sib.ch).
4C -----
4C EMBL: D82812; BA11576.1;
4C DB EMBL: D82815; BA11579.1;
4C InterPro: IPR01128; Cytochrome_P450.
4C Pfam: PF00067; P450_1.
4C PRINTS: PR00385; P450.
4C PROSITE: PS00086; CYTOCHROME_P450_1.
4C Oxidoreductase; Monooxygenase; Heme; NADP.
4C BINDING: 447 447 HEME (BY SIMILARITY).
4C SEQUENCE: 505 AA: 57973 MW: PAF970FA9B138151 CRG64;
4C -----
4C Query Match: 2.28; Score 7; DB 1; Length 505;
4C Best Local Similarity: 100.0%; Pred. No. 44;
4C Matches: 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
4C -----
4C 36 VLVALLY 42
4C |||||
4C 17 VLVALLY 23
4C -----
4C RESULT 18
4C TMA_LYCES
4C ID TMA_LYCES STANDARD; PRT; 527 AA.
4C AC Q22478;
4C 15-DEC-1998 (Rel. 36, Created)
4C 15-DEC-1998 (Rel. 37, Last sequence update)
4C 16-OCT-2001 (Rel. 40, Last annotation update)
4C Importin alpha subunit (tauoypharin alpha subunit) (KAP alpha).
4C Lycopersicon esculentum (Tomato).
4C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
4C Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
4C Asteridae; easterids I; Solanales; Solanaceae; Solanum.
4C NCBI_TaxID=4081;
4C [1]
4C SEQUENCE FROM N.A.
4C RA Kuehl T., Murrer L., Citovsky V., Catal Y.;
4C Submitted (JUN-1998) to the EMBL/GenBank/DBP databases
4C -1- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING
4C EITHER A SIMPLE OR BIPHARITE NLS MOTIF. PROMOTES DOCKING OF IMPORT
4C SUBSTRATES TO THE NUCLEAR ENVELOPE. SEEMS TO ACT AS A CYTOSOLIC
4C PROTEIN FOR BOTH SIMPLE AND BIPHARITE NLS MOTIFS (BY SIMILARITY).
4C -1- SUBUNIT: FORMS A COMPLEX WITH IMPORTIN BETA-1 SUBUNIT (BY
4C SIMILARITY).
4C -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
4C -1- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.
4C -1- SIMILARITY: CONTAINS 8 ARM REPEATS.
4C -----
4C This SWISS-PROT entry is copyright. It is produced through a collaboration
4C between the Swiss Institute of Bioinformatics and the EMBL outstation -
4C the European Bioinformatics Institute. There are no restrictions on its
4C use by non-profit institutions as long as its content is in no way
4C modified and this statement is not removed. Usage by and for commercial
4C entities requires a license agreement (See http://www.isb-sib.ch/announce/
4C or send an email to license@isb-sib.ch).
4C -----
4C EMBL: AF017252; AAC23722.1;
4C DB HSSP: G02421; 1BKS.
4C InterPro: IPR002525; Armadillo.
4C InterPro: IPR002522; ImportinA_B.
4C Pfam: PF00514; Armadillo_seq_8.
4C SMART: SM00185; ARM; 8.
4C PROSITE: PS0176; ARM_REPEAT; 5.
4C Transport; Protein transport; Repeat.
4C P-MAIN 12 51 ARM 1.
4C REPEAT 109 151 ARM 1.
4C REPEAT 152 196 ARM 2.

```

FT REPEAT 197 234 ARM 3.
 FT REPEAT 235 279 ARM 4.
 FT REPEAT 280 319 ARM 5.
 FT REPEAT 320 362 ARM 6.
 FT REPEAT 363 403 ARM 7.
 FT REPEAT 404 445 ARM 8.
 FT REPEAT 446 527 ARM 9.
 FT DOMAIN 446 527 ASP/GDU-RICH (ACIDIC).
 SO SEQUENCE 527 AA: 58605 MW: 443601641CFF4817 C6C64;
 Query Match 2.24; Score 7; DB 1; Length 527;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 LGICALV 24
 DB 198 LGICALV 204
 RESULT 19
 LAC1_PHLA STANDARD; PRT: 548 AA.
 ID LAC1_PHLA
 AC Q01679;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Laccase precursor (PF 1 to 3 2) (Ponzeoideae; oxygen oxidoreductase)
 DE (Urishiol oxidase) (Ligninolytic phenoloxidase).
 GN LAC.
 OS Phlebia radiata (White-rot fungus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Apophthorales; Corticiaceae; Phlebia.
 CX NCBI_TaxID=5308;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE
 RC STRAIN=ATCC 64658;
 RX MEDLINE=92065223; PubMed 1955850.
 RA Saloheimo M., Niku-Paavola M.L., Knowles T.K.;
 RT *Isolation and structural analysis of the laccase gene from the
 RT lignin-degrading fungus *Phlebia radiata*.
 RT J. Gen. Microbiol. 137:1547-1544(1991).
 CC
 CC -1- FUNCTION: MOST PROBABLY PLAYS AN IMPORTANT ROLE IN LIGNIN
 CC DEGRADATION
 CC -1- CATALYTIC ACTIVITY: 4 benzenediol + O(2) = 4 benzoquinone + 2
 CC H(2)O.
 CC -1- COFACTOR: BINDS 4 CU-IONS PER MOLECULE. THREE DISTINCT CU
 CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE
 CC 3 OR COUPLED BINUCLEAR.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES
 CC -1- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X52134; CAA36379.1;
 CC EMBL: A20705; CAA01513.1;
 CC PIR: S18746; S18746.
 CC HSSP: P37064; IASO.
 CC InterPro: IPR001117; Cu-oxidase.
 CC InterPro: IPR002355; Multicu_oxidase2.
 CC Pfam: PF00394; Cu-oxidase2.
 CC PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.
 CC PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
 CC Oxidoreductase; Signal; Copper; Metal-binding; Lignin degradation;
 CC Glycoprotein; Repeat.
 CC SIGNAL 1 21
 CC FT SIGNAL 22 548 LACCASE.
 CC FT DOMAIN 22 148 PLASTOCYANIN-LIKE 1.

FT DOMAIN 160 304 PLASTOCYANIN-LIKE 2.
 FT DOMAIN 373 496 PLASTOCYANIN-LIKE 3.
 FT METAL 87 85 COPPER (TYPE 2) (PROBABLE).
 FT METAL 87 85 COPPER (TYPE 3) (PROBABLE).
 FT METAL 130 130 COPPER (TYPE 4) (PROBABLE).
 FT METAL 132 132 COPPER (TYPE 3) (PROBABLE).
 FT METAL 132 132 COPPER (TYPE 3) (PROBABLE).
 FT METAL 421 421 COPPER (TYPE 2) (PROBABLE).
 FT METAL 423 423 COPPER (TYPE 3) (PROBABLE).
 FT METAL 423 423 COPPER (TYPE 3) (PROBABLE).
 FT METAL 474 474 COPPER (TYPE 1) (PROBABLE).
 FT METAL 474 474 COPPER (TYPE 1) (PROBABLE).
 FT METAL 475 475 COPPER (TYPE 3) (PROBABLE).
 FT METAL 479 479 COPPER (TYPE 1) (PROBABLE).
 FT METAL 484 484 COPPER (TYPE 1) (PROBABLE).
 FT CARBOHYD 75 75 N-LINKED (GLCNAC...) (POST-5
 FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POST-5
 FT CARBOHYD 402 402 N-LINKED (GLCNAC...) (POST-5
 FT CARBOHYD 457 457 N-LINKED (GLCNAC...) (POST-5
 SO SEQUENCE 548 AA: 58601 MW: 44567723656B2 C6C64;
 Query Match 2.24; Score 7; DB 1; Length 548;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;
 QY 308 DVDNTWI 314
 DB 275 DVDNTWI 281
 RESULT 20
 PRIZ_DROVI STANDARD; PRT: 584 AA.
 ID PRIZ_DROVI
 AC Q24760;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Friczled protein precursor (Friczled-1) (DZ1).
 GN FZ.
 OS Drosophila virilis (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Muscophora;
 OC Muscomorpha; Ephydroidea; Eusophyllidae; Drosophila.
 CX NCBI_TaxID=7244;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ovary;
 RX MEDLINE=96366421; PubMed=8770598;
 RA Jones K.H., Liu J., Adler P.N.;
 RT *Molecular analysis of EMS induced friczed mutations in
 RT melanaster*;
 RT Genetics 142:205-215(1996).
 CC
 CC -1- FUNCTION: Receptor for Wnt proteins. Most of Friczled
 CC are coupled to the beta-catenin canonical signaling
 CC leads to the activation of dishevelled proteins. It
 CC GSK-3 kinase, nuclear accumulation of beta-catenin
 CC of Wnt target genes. A second signaling pathway in
 CC calcium fluxes has been seen for some family member.
 CC not yet clear if it represents a distinct pathway
 CC integrated in the canonical pathway, as PKC seems to
 CC for Wnt mediated activation of GSK-3 kinase. Both
 CC to involve interactions with G-proteins. Regulated
 CC the cytoskeletons of epidermal cells to produce a
 CC of cuticular hairs and bristles.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (P...
 CC -1- DOMAIN: Lys-Thr-X-X-X-Tip motif is involved in the
 CC the Wnt/beta-catenin signaling pathway (by similarity
 CC similarity). The fz domain is involved in binding with a
 CC -1- SIMILARITY: BELONGS TO FAMILY FZ/SW/OF G PROTEIN
 CC RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 1 FRICZLED (FZ) DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a

between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: L4163; AAB38383.1; ALT_INT.
EMBL: L4340; NOT_ANNOTATED_CDS.
EMBL: L4341; NOT_ANNOTATED_CDS.
EMBL: L4342; NOT_ANNOTATED_CDS.
EMBL: L4343; NOT_ANNOTATED_CDS.
EMBL: L4344; NOT_ANNOTATED_CDS.
EMBL: FB90014841; DVI/NIZ.
InterPro: IPR000539; Fz_domain.
InterPro: IPR000832; GPCR_secretin.
Pfam: PF01534; Fz12led; I.
Pfam: PF01392; Fz; I.
PRINTS: PR00489; FRIZZLED.
SMART: SM00063; Fz; I.
PROSITE: PS50038; Fz; I.
PROSITE: PS50261; G_PROTEIN_RECEP_F2_1
Receptor, G-protein-coupled receptor, Transmembrane,
developmental protein; glycoprotein, signal.
SIGNAL.
CHAIN 1
DOMAIN 32 583
DOMAIN 32 251
TRANSMEM 252 272
DOMAIN 273 283
TRANSMEM 284 304
TRANSMEM 305 345
TRANSMEM 346 366
TRANSMEM 367 382
TRANSMEM 383 403
TRANSMEM 404 427
TRANSMEM 428 448
DOMAIN 449 473
TRANSMEM 474 494
TRANSMEM 495 534
TRANSMEM 535 555
DOMAIN 556 583
DOMAIN 51 169
SITE 558 563
SITE 581 583
CARBOHYD 70 70
CARBOHYD 170 170
SEQUENCE 583 AA; 65070 MW; DR1558PERPC30F5 CP064;
Query Match 2.28; Score 7; DB 1; Length 583.
test Local Similarity 100.08; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LGGLACL 18
IIIIII
DB 486 LALLACL 492

RESULT 21

PTMA_BACSU STANDARD: PRT; 610 AA.
AC P42956; P44435;
DE 01-NOV-1995 (Rel. 32, Created)
DE 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE PTS SYSTEM, manual-specific (IAB component) (ETIAR-M1) (Manual)
DE premase (IAB component) (Phosphotransferase enzyme II, ABC
DE component) (EC 2.7.1.69) (ETI-M1).
GN MTLA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX MBL_TaxID-1423;
RN [1]

SEQUENCE FROM N.A.
SIPAIN-168;
AC A-994 F., Fujita K., Sugawara I., Nakamura K., Yamane K.;
RL Submitted (SEP 1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SIPAIN-168;
RX MEDLINE 9712189; PubMed-8969502;
RY Yamane K., Kumano M., Kurita K.;
RT The 25 degrees to 36 degrees region of the Bacillus subtilis chromosome:
determination of the sequence of a 146 kb segment and identification
of 113 genes.*
FT Motophoretic 147-694/-695b(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC SIPAIN-168;
RX MEDLINE 9804403; PubMed-788377;
RY Kurita K., Matsuda N., Moszer I., Albertini A.M., Alloni G.,
Azevedo V., Portero M.G., Passieres P., Pololin A., Porcheret S.,
Koriss K., Kourisier L., Graus A., Braun M., Brignell S.C., Bron S.,
Brouillet S., Puschel G.V., Caldwell B., Capuano V., Carter N.M.,
Choi S.K., Cobanli J.J., Conterton J.F., Cummings N.J., Daniel R.A.,
Deinert J., Levine K.M., Fosterhelfer A., Ethlich S.D., Emerson P.T.,
Enrian K.D., Errington J., Fabre C., Ferrati E., Fougere D.,
Fritze C., Fujita M., Fujita Y., Funa S., Gallitz A., Galleron N.,
Ghim S.Y., Glaser P., Gottau A., Gollighly F.J., Grandi G.,
Guspi G., Guy R.J., Hagi K., Harech J., Harwood C.R., Heant A.,
Hilbert H., Hildebrandt S., Hosono S., Hulle M.F., Iraya M., Jones L.,
Jolis F., Karamata E., Katsura Y., Kiear Blanchard M., Klein C.,
Kobayashi Y., Koeber F., Katsushita G., Kioh S., Kumano M.,
Kurita K., Lapius A., Lathion S., Lauber J., Lazarevic V.,
Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
Medina N., Meliade R.P., Mizuno M., Mostl P., Nakai S., Noback M.,
Neeve T., O'Neill M., Ogawa K., Gollwa A., Oudega R., Park S.H.,
Parron V., Pohl T.M., Portetle D., Potwoll S., Prescott A.M.,
Prescan F., Pujic P., Purnelle B., Rapoport G., Ray M., Reynolds S.,
Rieger M., Rivolta C., Roeha F., Roche R., Rose M., Sadale Y.,
Sato T., Scanlan E., Schleich S., Schroeder R., Scroton F.,
Sekiguchi J., Sekowska A., Seur S.J., Sartor P., Shin R.S., Soldo B.,
Sorkin A., Tavoron F., Takagi T., Takahashi H., Takemaru K.,
Takuchi M., Tanakoshi A., Tanaka T., Tufstra P., Teyoni A.,
Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
Viari A., Wambit R., Wedler E., Wedler H., Weizenecker T.,
Winzels P., Wipat A., Yamamoto H., Yamane K., Yasunori K., Yata K.,
Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zanchin A.;
RT The complete genome sequence of the Gram-positive bacterium Bacillus
subtilis.*
PL Nature 396:249-256(1997).
CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOTOLYPHATE-DEPENDENT
SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS). A MAJOR CARBOHYDRATE ACTIVE
-!- TRANSPORT SYSTEM THE 11CD DOMAINS CONTAIN THE SUGAR BINDING SITE
AND THE TRANSMEMBRANE CHANNEL. THE 11A DOMAIN CONTAINS THE PRIMARY
PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-PHP); 11A TRANSFERS ITS
PHOSPHORYL GROUP TO THE 11B DOMAIN WHICH FINALLY TRANSFERS IT TO
THE SUGAR.
CC -!- CATALYTIC ACTIVITY: protein N-phosphotransferase + sugar -> protein
histidine + sugar phosphate.
CC -!- STRUCTURAL LOCATION: integral membrane protein
CC -!- SIMILARITY: CONTAINS 1 PTS ETIA DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PTS E1B DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PTS E1C DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR HSSP: P00550; 1A3A.
 DR Subtilisin; BGI1215; mLA.
 DR InterPro: IPR004718; PTS1IC_mLA.
 DR InterPro: IPR002178; PTS_EIIA_2.
 DR Pfam: PF00359; PTS_EIIA_2; 1.
 DR ProDom: PD001689; PTS_EIIA_2; 1.
 DR TricFAMS: TricF00851; mLA; 1.
 DR PROSITE: PS00372; PTS_EIIA_2; 1.
 KW Phosphotransferase system; Sugar transport; Transferase;
 KM Phosphotransferase system; Complete proteome;
 FT DOMAIN 1 339
 FT DOMAIN 340 495
 FT DOMAIN 495 610
 FT TRANSMEM 30 50
 FT TRANSMEM 54 74
 FT TRANSMEM 90 110
 FT TRANSMEM 141 161
 FT TRANSMEM 167 187
 FT TRANSMEM 189 209
 FT TRANSMEM 219 239
 FT TRANSMEM 251 271
 FT TRANSMEM 290 310
 FT TRANSMEM 319 339
 FT MOD_RES 262 262
 FT MOD_RES 547 547
 FT CONFLICT 134 135
 SV SEQUENCE 610 AA; 65426 MW; E446F36D50E5A907 CRC64;

Query Match 2.28; Score 7; DB 1; Length 610;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 33 LAGVIVA 39
 |||||
 DB 322 LAGVIVA 328

RESULT 22
 ID ABL1_MOUSE STANDARD; PRT: 708 AA.
 AC Q9QXJ1; 008642;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Amyloid beta A4 precursor protein-binding family B member 1 (Fe65
 DE protein).
 GN ABL1 OR FE65.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laktionova A.V., Ivanova N.B., Belyavsky A.V.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 498-674 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=97049965; PubMed=8894693;
 RA Bressler S.L., Gray M.D., Sopher B.L., Hu Q., Hearn M.G., Pham D.G.,
 RA Dinulos M.R., Pukich K., Sisodia S.S., Miller M.A., Distchev C.M.,
 RA Martin G.M.;
 RT "cDNA cloning and chromosome mapping of the human Fe65 gene;
 RT interaction of the conserved cytoplasmic domains of the human beta-
 RT amyloid precursor protein and its homologues with the mouse Fe65
 RT protein.";
 RL Hum. Mol. Genet. 5:1589-1598(1996).
 CC -1- FUNCTION: MAY MODULATE THE INTERNALIZATION OF BETA-AMYLOID
 CC PRECURSOR PROTEIN (BY SIMILARITY).
 CC -1- SUBUNIT: BINDS TO THE INTRACELLULAR DOMAIN OF THE BETA-AMYLOID
 CC PRECURSOR PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 PID DOMAINS.

CC CC This SWISS-Prot entry is copyright. It is produced through
 CC collaboration between the Swiss Institute of Bioinformatics and the
 CC the European Bioinformatics Institute. There are no restrictions
 CC use by non-profit institutions as long as its content and its
 CC modified and this statement is not removed. Usage by
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/ebis/oln/oln.html> or send an email to license@ebi.ac.uk).
 CC
 DR EMBL: AE206720; AAF20141.1;
 DR EMBL: L77665; AAB51603.1;
 DR MGD: MGI:107765; Appb1.
 DR InterPro: IPR000050; PID_domain.
 DR InterPro: IPR001202; WW_Fsp5_WWP.
 DR Pfam: PF00397; WW; 1.
 DR Pfam: PF00640; PID; 2.
 DR SMART: SM00452; PTB; 2.
 DR SMART: SM00456; WW; 1.
 DR PROSITE: PS01179; PID; 2.
 DR PROSITE: PS01159; WW_DOMAIN_1; 1.
 DR PROSITE: PS50020; WW_DOMAIN_2; 1.
 KW Repeat.
 FT DOMAIN 253 285
 FT DOMAIN 370 507
 FT DOMAIN 540 697
 FT DOMAIN 158 171
 SV SEQUENCE 708 AA; 77138 MW; EFB878176FB998F2 CRC64;

Query Match 2.28; Score 7; DB 1; Length 708;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 47 VPSSTLSQ 54
 |||||
 DB 3 VPSSTLSQ 9

RESULT 24
 ID ABL1_HUMAN STANDARD; PRT: 710 AA.
 AC Q00213;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Amyloid beta A4 precursor protein-binding family B mem-
 DE protein).
 GN ABL1 OR FE65.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute-
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; H.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Laktionova A.V., Ivanova N.B., Belyavsky A.V.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 498-674 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97013428; PubMed 9799084;
 RA Hu Q., Kukull W.A., Bressler S.L., Gray M.D., Cam J.A.,
 RA Martin G.M., Jeeb S.S.;
 RT "The human Fe65 gene: genomic structure and an intronic
 RT polymorphism associated with sporadic dementia of the A
 RT type.";
 RL Hum. Genet. 103:295-303(1998).


```

CC -1- FUNCTION: CATALYZES THE CROSS-LINKING OF PROTEINS AND THE
CC CONJUGATION OF POLYAMINES TO PROTEINS. CONTRIBUTES TO THE
CC FORMATION OF THE CORNIFIED CELL ENVELOPE OF KERATINOCYTES.
CC -1- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N(5)-
CC alkylglutamine + NH(3).
CC -1- COFACTOR: Binds 1 calcium ion per subunit (by similarity).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE)
CC AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN FORESKIN KERATINOCYTES.
CC -1- SIMILARITY: BELONGS TO THE TRANSGLUTAMINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF05960; AAC02978.1; -
DR EMBL: AF05961; AAC02979.1; -
DR EMBL: AF206510; AAF23981.1; -
DR EMBL: AF206502; AAF23981.1; JOINED.
DR EMBL: AF206503; AAF23981.1; JOINED.
DR EMBL: AF206504; AAF23981.1; JOINED.
DR EMBL: AF206505; AAF23981.1; JOINED.
DR EMBL: AF206506; AAF23981.1; JOINED.
DR EMBL: AF206507; AAF23981.1; JOINED.
DR EMBL: AF206508; AAF23981.1; JOINED.
DR EMBL: AF206509; AAF23981.1; JOINED.
DR HSSP: P00488; IQK.
DR GENE: HGNC:11781; TGM5.
DR MIM: 603805; -
DR InterPro: IPR001102; GlutraminG.
DR InterPro: IPR002931; Transglutaminase_like.
DR Pfam: Pf00868; Transglutamin_N; 1.
DR Pfam: Pf01841; Transglut_core; 2.
DR SMART: SM00460; TGC; 1.
DR PROSITE: PS00547; TRANSGLUTAMINASES; 1.
KW Transglutaminase; Acyltransferase; Calcium-binding; Polymorphism;
KW Alternative splicing.
FT ACT_SITE 278 278 BY SIMILARITY.
FT ACT_SITE 337 337 BY SIMILARITY.
FT ACT_SITE 360 360 BY SIMILARITY.
FT CA_BIND 400 400 BY SIMILARITY.
FT CA_BIND 402 402 BY SIMILARITY.
FT CA_BIND 448 448 BY SIMILARITY.
FT CA_BIND 453 453 BY SIMILARITY.
FT VARSPIC 64 145 MISSING (IN SHORT ISOFORM).
FT VARIANT 67 67 P -> S.
FT VARIANT 352 352 /FTID-VAR_013248.
FT VARIANT 352 352 A -> G.
FT SEQUENCE 720 AA: 80777 MW: 9656884.848 kDa
FT SQU 720 AA: 80777 MW: 9656884.848 kDa
Query Match 2.28; Score 7; DB 1; Length 720;
Best Local Similarity 100.0%; Pred No 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DE 1/CR3/p150,95 beta-subunit) (cattle) (Complement receptor
DE subunit).
DE ITGB2 OR CD18.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Eute-
OC mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP MEDLINE:93290287; PubMed 1451021;
RA Shuster D.E., Bosworth R.T., Kehrl M.F. Jr.;
RT "Sequence of the bovine CD18-encoding cDNA: comparison w
RT and murine glycoproteins."
PL Gene 114:267-271(1992).
PP [2]
PP VARIANT LAD GLY-128.
PP MEDLINE:93028437; PubMed-1484045;
RA Shuster D.E., Kehrl M.F. Jr., Ackermann M.R., Gilbert P.
RT "Identification and prevalence of a genetic defect that c
RT leukocyte adhesion deficiency in Holstein cattle."
PL Proc Natl Acad Sci U S A 89:9245-9249(1992).
CC 1
CC FUNCTION: INTEGRIN ALPHA-1/BEI-2 IS A RECEPTOR FOR
CC ICAM3 AND ICAM4. INTERACTS ALPHA-M/BETA-2 AND ALPHA-X
CC RECEPTORS FOR THE IC3H FRAGMENT OF THE THIRD COMPLEMENT
CC AND FOR FIBRIN-GEN. INTERIN ALPHA-X/BETA-2 RECOGNIZE
CC SEQUENCE G-P-R IN FIBRINOGEN ALPHA-CHAIN. INTEGRIN A1
CC RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHA
CC ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FACTOR X. INTER
CC D/BETA-2 IS A RECEPTOR FOR ICAM3 AND VCAM1.
CC -1- SHUNT: HEPHODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC ASSOCIATES WITH EITHER ALPHA-L, ALPHA-M, ALPHA-X OR
CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
CC -1- DISEASE: DEFECTS IN ITGA2 ARE THE CAUSE OF LEUKOCYTE
CC DEFICIENCY (LAD). THE MUTATION CAUSING LAD (GLY-128)
CC AMONG HOLSTEIN CATTLE THROUGHOUT THE WORLD, PLACING
CC AMONG THE MOST COMMON GENETIC DISEASES KNOWN IN ANIMA
CC AGRICULTURE. ALL CATTLE WITH THE MUTANT ALLELE ARE PR
CC BULL. WHO THROUGH THE USE OF ARTIFICIAL INSEMINATION
CC CALVES IN THE 1950S AND 1960S.
CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMIL
CC -1- SIMILARITY: CONTAINS 1 WFA-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced throug
CC between the Swiss Institute of Bioinformatics and the
CC the European Bioinformatics Institute. There are no res
CC use by non-profit institutions as long as its content
CC modified and this statement is not removed. Usage by a
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M81233; AAA30438.1; -
DR EMBL: J01121; J01121.
DR HSSP: P05106; J0V2.
DR InterPro: IPR000561; EGF-1-like.
DR InterPro: IPR002369; Integrin_B.
DR InterPro: IPR001169; Integrin_beta_C.
DR InterPro: IPR003659; Plectin-like.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF00362; Integrin_B; 1.
DR PRINTS: PR01186; INTEGRINB.
DR PRODOM: PD001811; Integrin_B; 1.
DR SMART: SM00187; INB; 1.
DR SMART: SM00423; PST; 1.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS000243; INTEGRIN_BETA_3.
DR PROSITE: PS000022; EGF_1; EGF-WR-2.
DR PROSITE: PS01186; EGF_2; UNKNOWN-2.
KW Integrin; Cell adhesion; Receptor; Transmembrane; Glycop
KW Repeat; Signal; Disease mutation.
FT CHAIN 1 22 BY SIMILARITY.
FT CHAIN 23 769 INTEGRIN_BETA-2.
FT DOMAIN 23 700 EXTRACELLULAR (POTENTIAL).

```

FT	TRANSMEM	701	723	POTENTIAL.
FT	DOMAIN	724	769	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	124	363	WFA-LIKE.
FT	DOMAIN	449	617	4 CYSTEINE-RICH TANDEM REPEATS.
FT	REPEAT	449	496	1.
FT	REPEAT	497	540	11.
FT	REPEAT	541	581	11.
FT	REPEAT	582	617	11.
FT	DISULFID	25	447	BY SIMILARITY.
FT	DISULFID	33	43	BY SIMILARITY.
FT	DISULFID	36	73	BY SIMILARITY.
FT	DISULFID	46	62	BY SIMILARITY.
FT	DISULFID	191	198	BY SIMILARITY.
FT	DISULFID	246	286	BY SIMILARITY.
FT	DISULFID	386	400	BY SIMILARITY.
FT	DISULFID	420	662	BY SIMILARITY.
FT	DISULFID	445	449	BY SIMILARITY.
FT	DISULFID	459	470	BY SIMILARITY.
FT	DISULFID	467	506	BY SIMILARITY.
FT	DISULFID	472	481	BY SIMILARITY.
FT	DISULFID	483	497	BY SIMILARITY.
FT	DISULFID	512	517	BY SIMILARITY.
FT	DISULFID	514	549	BY SIMILARITY.
FT	DISULFID	519	534	BY SIMILARITY.
FT	DISULFID	536	541	BY SIMILARITY.
FT	DISULFID	557	562	BY SIMILARITY.
FT	DISULFID	559	590	BY SIMILARITY.
FT	DISULFID	564	573	BY SIMILARITY.
FT	DISULFID	575	582	BY SIMILARITY.
FT	DISULFID	596	601	BY SIMILARITY.
FT	DISULFID	598	643	BY SIMILARITY.
FT	DISULFID	603	612	BY SIMILARITY.
FT	DISULFID	615	618	BY SIMILARITY.
FT	DISULFID	622	631	BY SIMILARITY.
FT	DISULFID	628	695	BY SIMILARITY.
FT	DISULFID	647	670	BY SIMILARITY.
FT	SITE	397	399	CELL ATTACHMENT SITE (POTENTIAL).
FT	CARBOHYD	50	50	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	116	116	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	254	254	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	501	501	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	642	642	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	128	128	N > G (IN LAD).
NO	SEQUENCE	769 AA:	844aa MW: 500404690pI: 4.76	
Quantity Match 2.2% Score 71 DB 1; Length 769;				
Host Local Similarity 100.0% Pred. No. 65;				
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Q7	71 AVGISE 77			
Id	441 AVGISE 347			

PT	"Cloning of the beta subunit of the leukocyte adhesion proteins:
PT	homology to an extracellular matrix receptor defines a novel
RT	supergene family."
RL	Cell 48:681-690(1987).
RL	[2]
RN	SEQUENCE FROM N.A.
RX	MEDLINE=92077153; PubMed=1684838;
RA	Welleman J.R., Wells C.E., Wright A.H., Clark P.A., Law S.K.A.;
ET	"The gene organization of the human beta 2 integrin subunit (CD18)."
RL	PNAS Lett. 294:97-103(1991).
RL	[3]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=20289799; PubMed=10840953;
RA	Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
PA	Park H.-S., Toyoda A., Ishii K., Yokoi Y., Choi D.-K., Soeda E.,
PA	Onji M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Bulley A.,
PA	Menzel U., Holthar J., Kempt K., Lehmann P., Patterson D.,
PA	Reichwald K., Pump A., Schillhabel M., Schudy A., Zimmermann W.,
PA	Pasenthal A., Kudoh T., Shibuya K., Kawasaki K., Asakawa S.,
PA	Shimizu A., Sakaki T., Narasawa K., Mitsuyama S., Anticardias S.E.,
PA	Shimizu S., Shimizu N., Morisaki G., Hornischer K., Brandt P.,
PA	Shattil M., Shattil O., Gsaller A., Reiblich J., Kauz G., Blocker H.,
PA	Kramer J., Beck A., Klages S., Hennig S., Kieselmann T., Dagand R.,
PA	Wohrner S., Korym K., Gardiner K., Nitzel D., Francis F.,
PA	Lehrach H., Reinhardt P., Vasspe M.-L.;
PL	"The DNA sequence of human chromosome 21.";
PL	Nature 405:411-419(2000).
PL	[4]
RP	SEQUENCE OF 9-769 FROM N.A. AND PARTIAL SEQUENCE.
RC	TISSUE=SpLenn;
RX	MEDLINE=87244525; PubMed=2954816;
RA	Law S.K.A., Gaudon J., Hildreth J.E., Wells C.E., Willis A.C.,
RA	Wong A.J.;
ET	"The primary structure of the beta-subunit of the cell surface
ET	adhesion glycoproteins LFA-1, CR3 and p150,95 and its relationship to
ET	the fibronectin receptor."
RL	EMBO J. 6:915-919(1987).
RL	[5]
RN	SEQUENCE OF 347-365 FROM N.A. AND VARIANTS LAD1 SEP-351 AND LAD-586.
RX	MEDLINE=92137641; PubMed=1346613;
RA	Nelson C., Radd H., Arnout M.A.;
ET	"Genetic cause of leukocyte adhesion molecule deficiency: Abnormal
ET	splicing and a missense mutation in a conserved region of CD18 impair
ET	cell surface expression of beta 2 integrins."
RL	J. Biol. Chem. 267:3351-3357(1992).
RL	[6]
RP	VARIANTS LAD1 PRO-149 AND ARG-169.
RX	MEDLINE=90234666; PubMed=1944220;
RA	Wardlaw A.J., Hibbs M.L., Stacker S.A., Springer T.A.;
ET	"Persistent mutations in two patients with leukocyte adhesion
ET	deficiency and their functional correlates."
RL	J. Exp. Med. 172:335-345(1990).
RL	[7]
RP	VARIANT LAD1 ARG-169.
RX	MEDLINE=92324303; PubMed=1352501;
RA	Corbi A., Varga A., Orsa A., Rodriguez M.C.G., Fontan G.,
RA	Sanchez-Madrid F.;
ET	"Molecular basis for a severe case of leukocyte adhesion deficiency."
RL	Eur. J. Immunol. 22:1877-1881(1992).
RL	[8]
RP	VARIANTS LAD1 THR-196 AND CYS-593.
RX	MEDLINE=90187099; PubMed=1968911;
RA	Arizumi M.A., Chan N., Gupta S.K., Jacob J.C., Fathallah D.M.;
ET	"Point mutations impairing cell surface expression of the common beta
ET	subunit (CD18) in a patient with leukocyte adhesion molecule
ET	(LAD-CAM) deficiency."
RL	J. Clin. Invest. 85:977-981(1990).
RL	[9]
RP	VARIANT LAD1 LEU-178.
RX	MEDLINE=92184805; PubMed=1475332;
RA	Back L.L., Kwok W.W., Hickey D.D.;
ET	"Identification of two molecular defects in a child with leukocyte
ET	adhesion deficiency."

RL J Biol. Chem. 267:5482-5487(1992).
 RN [10]
 RP VARIANT LAD1 ASN-128.
 RX MEDLINE=92272746; PubMed=1590804.
 RA Matsura S., Kishi F., Tsukahara M., Nunoi H., Matsuda I.,
 KA Kobayashi K., Kajii T.
 RT "Leukocyte adhesion deficiency: identification of novel mutations in
 RT two Japanese patients with a severe form."
 RL Biochem. Biophys. Res. Commun. 184:1450-1457(1992).
 RN [11]
 RP VARIANT LAD1 SER-284.
 RX MEDLINE=9312347; PubMed=7686755;
 RA Hack L.A., Kerkering M., Baker D., Hauser I.R., Embree L.J.,
 KA Hickey D.D.
 RT "A point mutation associated with leukocyte adhesion deficiency type
 RT 1 of moderate severity."
 RL Biochem. Biophys. Res. Commun. 193:912-918(1993).
 RN [12]
 RP VARIANT LAD1 PRO-138 AND ARG-273.
 RX MEDLINE=94102253; PubMed=9884334;
 KA Hoeg N., Stewart M.P., Smith S.L., Newton P., Shaw T.M., Jew S.K.A.,
 KA Klein N.
 RT "A novel leukocyte adhesion deficiency caused by expressed but
 RT nonfunctional beta2 integrins Mac-1 and LFA-1."
 RL J. Clin. Invest. 103:97-106(1999).
 CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR ICAM1, ICAM2,
 CC ICAM3 AND ICAM4. INTEGRINS ALPHA-M/BETA-2 AND ALPHA-X/BETA-2 ARE
 CC RECEPTORS FOR THE IC3H FRAGMENT OF THE THIRD COMPLEMENT COMPONENT
 CC AND FOR FIBRINOGEN. INTEGRIN ALPHA-X/BETA-2 RECOGNIZES THE
 CC SEQUENCE G-P-R IN FIBRINOGEN ALPHA-CHAIN. INTEGRIN ALPHA-M/BETA-2
 CC RECOGNIZES P1 AND P2 PEPTIDES OF FIBRINOGEN GAMMA CHAIN. INTEGRIN
 CC ALPHA-M/BETA-2 IS ALSO A RECEPTOR FOR FACTOR X. INTEGRIN ALPHA-
 CC D/BETA-2 IS A RECEPTOR FOR ICAM3 AND VCAM1.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-2
 CC ASSOCIATES WITH EITHER ALPHA-L, ALPHA-M, ALPHA-X OR ALPHA-D.
 CC -1- SURCELLULAR LOCATION: Type 1 membrane protein.
 CC -1- DISEASE: DEFECTS IN ITGB2 ARE THE CAUSE OF LEUKOCYTE ADHESION
 CC DEFICIENCY, TYPE I (LAD1). LAD1 PATIENTS HAVE RECURRENT BACTERIAL
 CC INFECTIONS AND THEIR LEUKOCYTES ARE DEFICIENT IN A WIDE RANGE OF
 CC ADHESION-DEPENDENT FUNCTIONS.
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 VWA-LIKE DOMAIN.
 CC -1- DATABASE: NAME-PROV: NITE-CD guide cd18 entry;
 CC WWW=HTTP://WWW.NITE.NIH.GOV/PROV/CD/CD18.HTM".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y00057; CAA68266.1; -;
 DR EMBL: X64072; CAA45427.1; JOINED.
 DR EMBL: X64074; CAA45427.1; JOINED.
 DR EMBL: X64075; CAA45427.1; JOINED.
 DR EMBL: X64076; CAA45427.1; JOINED.
 DR EMBL: X64077; CAA45427.1; JOINED.
 DR EMBL: X64078; CAA45427.1; JOINED.
 DR EMBL: X64079; CAA45427.1; JOINED.
 DR EMBL: X64080; CAA45427.1; JOINED.
 DR EMBL: X64081; CAA45427.1; JOINED.
 DR EMBL: X64082; CAA45427.1; JOINED.
 DR EMBL: X64083; CAA45427.1; JOINED.
 DR EMBL: X64924; CAA45427.1; JOINED.
 DR EMBL: X64925; CAA45427.1; JOINED.
 DR EMBL: X64926; CAA45427.1; JOINED.
 DR EMBL: A1163300; CAB90553.1; -;
 DR EMBL: M15195; AAA59490.1; -;
 DR EMBL: S81234; AAB21404.1; -;
 DR PIR: A25967; LJDULM.

DR PIR: A25965; A25965.
 DR HSSP: P05106; 1JV2.
 DR Genew: HGNC:6155; ITGB2.
 DR MIM: 600065; -;
 DR MIM: 116920; -;
 DR InterPro: IPR002864; PPF-11ko.
 DR InterPro: IPR001169; Integrin_beta_2.
 DR InterPro: IPR003659; Plexin-11ko.
 DR InterPro: IPR002035; VWF_A.
 DR Pfam: PF00362; Integrin_B_1.
 DR PRINTS: PD01186; INTEGRIN.
 DR ProDom: PD001811; Integrin_beta_1.
 DR SMART: SM00187; INH_1.
 DR SMART: SM00424; PSI_1.
 DR SMART: SM00327; VWA_1.
 DR PROSITE: PS00243; INTEGRIN_BETA_2.
 DR PROSITE: PS00222; EGF_1; CRK_RNL_2.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_4.
 DR Prosite: Cell adhesion, Receptor, Transmembrane Glycyl
 KW Repeat, Signal, Disease mutation.
 FT SIGNAL 1 22
 FT CHAIN 23 769 INTEGRIN BETA-2
 FT DOMAIN 23 700 EXTRACELLULAR (POTENTIAL)
 FT TRANSM 701 723 POTENTIAL
 FT DOMAIN 724 769 CYTOPLASMIC (POTENTIAL)
 FT DOMAIN 124 363 VWA-LIKE
 FT REPEAT 449 617 4 CYSTEINE-RICH TANDEM PEP
 FT REPEAT 497 540 1
 FT REPEAT 541 581 11
 FT REPEAT 582 617 11
 FT DISULFD 25 447 IV
 FT DISULFD 33 43 BY SIMILARITY
 FT DISULFD 36 73 BY SIMILARITY
 FT DISULFD 46 62 BY SIMILARITY
 FT DISULFD 191 198 BY SIMILARITY

Query Match 2.24; Score 7; DB: 1; Length 75
 Best Local Similarity 100.0%; Prod. No: 69;
 Matches 7; Conservative 0; Mismatches 0; Indel 0

QY 71 ANGELSE 77
 DB 341 ANGELSE 347
 ITR2_PIG
 ITR2_PIG STANDARD; EPI: 769 AA.
 AC P53714;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 41, Last annotation update)
 DE Integrin beta-2 precursor (cyl surface adhesion glycosylated)
 DE I/CR3/1156.95 beta-subunit (cyl) (Complement receptor subunit)
 DE subunit.
 GN ITGB2 OR CD18.
 OS Sus scrofa (Pig).
 OC Mammalia, Metazoa, Chordata, Vertebrata, Sui
 OC Mammalia, Theria, Cetartiodactyla, Suidae, Sui
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lee J.K., Schrock L.B., Rothbard M.S.
 RT "Molecular cloning and characterization of the porcine
 RT adhesion molecule."
 RL J. Biol. Chem. 267:2222-2230(1992).
 CC -1- FUNCTION: INTEGRIN ALPHA-L/BETA-2 IS A RECEPTOR FOR
 CC ICAM3 AND ICAM4. INTEGRINS ALPHA-M/BETA-2 AND ALPHA-X/BETA-2
 CC RECEPTORS FOR THE IC3H FRAGMENT OF THE THIRD COMPLEMENT
 CC AND FOR FIBRINOGEN. INTEGRIN ALPHA-X/BETA-2 RECOGNIZES THE
 CC SEQUENCE G-P-R IN FIBRINOGEN ALPHA-CHAIN. INTEGRIN

CC RECOGNIZES ONE OR MORE DOMAINS WITHIN THE ALTERNATIVELY SPLICED
 CC CS-1 REGION OF FIBRONECTIN. INTEGRIN ALPHA-4/BETA-7 IS ALSO A
 CC RECEPTOR FOR MADCAM1 AND VCAM1. IT RECOGNIZES THE SEQUENCE L-D-T
 CC IN MADCAM1. INTEGRIN ALPHA-E/BETA-7 (HML-1) IS A RECEPTOR FOR E-
 CC CADHERIN.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-7
 CC ASSOCIATES WITH EITHER ALPHA-4 OR ALPHA-E.
 CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF LEUKOCYTE LINES.
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 VWA-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright © 1999 by the EMBL
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL: S80335; AAB21332.1; -;
 DR EMBL: M62880; AAA59185.1; -;
 DR EMBL: M68892; AAA59184.1; -;
 DR EMBL: S49378; AAB21688.1; -;
 DR EMBL: S49364; AAB21688.1; JOINED.
 DR EMBL: S49365; AAB21688.1; JOINED.
 DR EMBL: S49366; AAB21688.1; JOINED.
 DR EMBL: S49367; AAB21688.1; JOINED.
 DR EMBL: S49368; AAB21688.1; JOINED.
 DR EMBL: S49369; AAB21688.1; JOINED.
 DR EMBL: S49370; AAB21688.1; JOINED.
 DR EMBL: S49371; AAB21688.1; JOINED.
 DR EMBL: S49373; AAB21688.1; JOINED.
 DR EMBL: S49374; AAB21688.1; JOINED.
 DR EMBL: S49375; AAB21688.1; JOINED.
 DR EMBL: S49377; AAB21688.1; JOINED.
 DR EMBL: L23823; AAA36118.1; -;
 DR EMBL: L23810; AAA36118.1; JOINED.
 DR EMBL: L23811; AAA36118.1; JOINED.
 DR EMBL: L23812; AAA36118.1; JOINED.
 DR EMBL: L23813; AAA36118.1; JOINED.
 DR EMBL: L23814; AAA36118.1; JOINED.
 DR EMBL: L23815; AAA36118.1; JOINED.
 DR EMBL: L23816; AAA36118.1; JOINED.
 DR EMBL: L23817; AAA36118.1; JOINED.
 DR EMBL: L23818; AAA36118.1; JOINED.
 DR EMBL: L23819; AAA36118.1; JOINED.
 DR EMBL: L23820; AAA36118.1; JOINED.
 DR EMBL: L23822; AAA36118.1; JOINED.
 DR PIR: A40526; A40526.
 DR HSSP: P05106; IJV2.
 DR Gene: HGNC:6162; ITGB7.
 DR MIM: 147559; -;
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR002369; Integrin_B.
 DR InterPro: IPR001169; Integrin_beta_C.
 DR InterPro: IPR003659; Plexin-like.
 DR InterPro: IPR002035; VWF_A.
 DR Pfam: PF00362; Integrin_B; 1.
 DR PRINTS: PR01186; INTEGRINB.
 DR ProDom: PD001811; Integrin_B; 1.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00187; INH; 1.
 DR SMART: SM00423; PSI; 1.
 DR PROSITE: PS00327; VWA; 1.
 DR PROSITE: PS00243; INTEGRIN_BETA_3.
 DR PROSITE: PS00022; EGF_1; UNKNOWN_4.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR Integrin: Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KW Repeat: Signal; Phosphorylation; Alternative splicing.
 FT SIGNAL
 FT 1
 FT 19

FT	CHAIN	20	798	INTEGRIN BETA-7.
FT	DOMAIN	20	723	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	724	746	POTENTIAL.
FT	DOMAIN	747	798	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	150	389	VWA-LIKE.
FT	DOMAIN	478	640	4 CYSTEINE-RICH TANDEN REPE
FT	REPEAT	478	526	1.
FT	REPEAT	527	565	11.
FT	REPEAT	566	604	11.
FT	REPEAT	605	640	11.
FT	DISULFID	45	476	BY SIMILARITY.
FT	DISULFID	51	61	BY SIMILARITY.
FT	DISULFID	54	91	BY SIMILARITY.
FT	DISULFID	64	80	BY SIMILARITY.
FT	DISULFID	216	223	BY SIMILARITY.
FT	DISULFID	271	311	BY SIMILARITY.
FT	DISULFID	412	428	BY SIMILARITY.
FT	DISULFID	448	688	BY SIMILARITY.
FT	DISULFID	474	478	BY SIMILARITY.
FT	DISULFID	488	500	BY SIMILARITY.
FT	DISULFID	497	537	BY SIMILARITY.
FT	DISULFID	502	511	BY SIMILARITY.
FT	DISULFID	513	527	BY SIMILARITY.
FT	DISULFID	543	548	BY SIMILARITY.
FT	DISULFID	545	574	BY SIMILARITY.
FT	DISULFID	550	559	BY SIMILARITY.
FT	DISULFID	561	566	BY SIMILARITY.
FT	DISULFID	580	585	BY SIMILARITY.
FT	DISULFID	582	613	BY SIMILARITY.
FT	DISULFID	587	596	BY SIMILARITY.
FT	DISULFID	598	605	BY SIMILARITY.
FT	DISULFID	619	624	BY SIMILARITY.
FT	DISULFID	621	666	BY SIMILARITY.
FT	DISULFID	626	635	BY SIMILARITY.
FT	DISULFID	638	641	BY SIMILARITY.
FT	DISULFID	645	654	BY SIMILARITY.
FT	CARBOHYD	279	279	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	434	434	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	477	477	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	531	531	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	590	590	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	665	665	N-LINKED (GLCNAC...) (P...)
FT	CARBOHYD	674	674	N-LINKED (GLCNAC...) (P...)
FT	MOD_RES	778	778	PHOSPHORYLATION (BY TYR-KIN...)
FT	VARSPLIC	501	648	MISSING (IN SHORT ISOFORM)
FT	MUTAGEN	159	159	D-EA: LOSS OF INTEGRIN ALI...
FT				BINDING TO E-CADHERIN AND
FT				ALPHA-4/BETA-7 BINDING TO E
FT	SEQUENCE	798 AA; 8650 MW; 1382750E9992785 CR64;		

Query Match 2.2%; Score 7; DB 1; Length 7;
 Best Local Similarity 100.0%; Prod. No. 67;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

Oy 71 AVGELSE 77
 Db 367 AVGLISE 374

RESULT 30
 SCYL_YEAST
 ID SCYL_YEAST STANDARD; (P): 804 AA.
 AC P51009;
 DT 01-OCT-1996 (rel. 34, Citedrel)
 DT 01-OCT-1996 (rel. 34, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE SCYL protein.
 DE SCYL OR YGL083W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycetes.
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

```

CX NCHL_TaxID=4932;
FN [1]
RP SEQUENCE FROM N.A.
KA Petersch M., Gallwitz D.;
KL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases
KN [2]
RP SEQUENCE FROM N.A.
KA STRAIN=5286c;
RX MEDLINE-97435481; PubMed 9290212;
RA Rieger M., Ruckelshaus M., Schaefer M., Mueller-Auer S.;
RT Sequence analysis of 203 Kinasases from Saccharomyces cerevisiae
RL Yeast 13:1077-1090(1997)
CX - SIMILARITY: SOME, TO THE CATALYTIC DOMAIN OF PROTEIN KINASES
C2 This SWISS-Prot entry is copyright. It is produced through a collaboration
C2 between the Swiss Institute of Bioinformatics and the EMBL outstation -
C2 the European Bioinformatics Institute. There are no restrictions on its
C2 use by non-profit institutions as long as its content is in no way
C2 modified and this statement is not removed. Usage by and for commercial
C2 entities requires a license agreement (See http://www.isb-sib.ch/announcement/
C2 or send an email to license@sib.ch)
CX -----
D3 EMBL: X97305; CAA65975.1;
D3 EMBL: Z72605; CAA96788.1;
D3 SCD: S0003051; SCY1
D3 InterPro: IPR000719; Euk.pkinase.
D3 Pfam: PF000069; Pkinase.1
D3 PROSITE: PS50011; PROTEIN_KINASE_DOM;
F1 DOMAIN 1 324 PROTEIN KINASE-LIKE.
F1 DOMAIN 36 43 POLY-SER.
SQ SEQUENCE 804 AA; 91000 MW; 457A81DDC7AB808E CRC64;

Query Match 2.2%; Score 7; DB 1; Length 804;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 234 SEQUENCE 242
ID [1]
DB 134 EGNFID 140

RESULT 31
YBBP_ECOLI STANDARD; PRT; 804 AA.
AC P77504;
D1 01-NOV-1997 (Rel. 35, Created)
D1 01-NOV-1997 (Rel. 35, Last sequence update)
D1 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybbp.
GN YBBP OF H0406.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
CX NCBI_TaxID=562;
[1]
RP SEQUENCE FROM N.A.
KA STRAIN=K12 / MG1655;
RX MEDLINE-97426617; PubMed 9278503;
KA Blattner F.R., Plunkett G. III, Bloch C.A., Bernal N.T., Holman V.,
KA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.W.,
KA Gregor T., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose T.J.,
KA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
[2]
RP SEQUENCE FROM N.A.
KA Roberts R., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
KA Duran M., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi D.,
KA Lew H., Lin D., Nemach A., Oetner P., Schramm S., Davis P.W.;
KL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CX - SIMILAR LOCATIONS: Integral membrane protein (Potential)

```

```

CC This SWISS-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib.ch)
CX -----
D3 EMBL: AE000155; AAC73598.1;
D3 EMBL: D82544; AAA0250.1;
D3 EcoGene: EG03263; ybbp.
D3 InterPro: IPR004838; DUF214.
D3 Pfam: PF02687; DUF214; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 24 263 POTENTIAL.
FT TRANSMEM 301 321 POTENTIAL.
FT TRANSMEM 334 353 POTENTIAL.
FT TRANSMEM 381 401 POTENTIAL.
FT TRANSMEM 403 423 POTENTIAL.
FT TRANSMEM 453 473 POTENTIAL.
FT TRANSMEM 680 700 POTENTIAL.
FT TRANSMEM 734 754 POTENTIAL.
FT TRANSMEM 769 789 POTENTIAL.
SQ SEQUENCE 804 AA; 89332 MW; 3475b8140b9207a CRC64;

Query Match 2.2%; Score 7; DB 1; Length 804;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 28 ISFMIL 34
ID [1]
DB 462 ISFMIL 468

RESULT 32
ITB7_MOUSE STAN/ABL; PRT; 806 AA.
AC P26011; O64656;
D1 01-MAY-1992 (Rel. 22, Created)
D1 01-JUL-1993 (Rel. 26, Last sequence update)
D1 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integrin beta-7 precursor (Integrin beta-P) (M290 IEL antigen).
GN ITB7.
OS Mus musculus (Mouse).
OC Eukaryota; Mammalia; Chordata; Craciata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE-92390425; PubMed 1518854;
FA Hu M.C.T., Orrego P.T., Weissman I.L., Holzman B.;
RT "Cloning and expression of mouse integrin beta 7 (beta 7): a
RT functional role in Peyer's patch-specific lymphocyte homing."
PL Proc. Natl. Acad. Sci. U.S.A. 89:8754-8758(1992).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE-92218384; PubMed 1579978;
FA Yuan Q., Jiang W. M., Leung K., Hollander D., Watson J.D.,
PA Kristiansen G.W.;
RT "Cloning of the mouse integrin beta 7 subunit."
PL J Biol Chem 267:7345-7349(1992).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE-92388644; PubMed 1341400;
PA Garish M.F., Bell A.F., Smith T.J., Ducharme L.A., Wang R.K.,
RA Weis J.H.;
RT "Expression of murine beta 7, alpha 4, and beta 1 integrin genes by
RT rodent mast cells."
PL J Immunol 149:1964-1972(1992).
[4]
RP SEQUENCE OF 1-90 FROM N.A., AND SEQUENCE OF 20-32.

```


RX MEDLINE-91248239, PubMed-1719115,
 RA Yuan Q., Jlang W.-M., Hollander D., Leung E., Watson J.D.,
 RA Katsansen G.W.:
 RT "Identity between the novel integrin beta 7 subunit and an antigen
 RT found highly expressed on intra-epithelial lymphocytes in the small
 RT intestine."
 RL Biochem. Biophys. Res. Commun. 176:1443-1449(1991).
 RN [5]
 RP SEQUENCE OF 1-67 FROM N.A.
 RA MEDLINE-93305607; PubMed-8318458;
 RA Leung E., Mead P.E., Yuan Q., Jlang W.M., Watson J.D.,
 RA Katsansen G.W.:
 RT "The mouse beta 7 integrin gene promoter: transcriptional regulation
 RT of the leukocyte integrin LPM-1 and M290."
 RL Immunol. 5:551-558(1993).
 CC -1- FUNCTION: INTEGRIN ALPHA-4/BETA-7 (PEYER'S PATCHES-SPECIFIC HOMING
 CC RECEPTOR LPM-1) IS INVOLVED IN ADHESIVE INTERACTIONS OF
 CC LEUKOCYTES. IT IS A RECEPTOR FOR FIBRONECTIN AND RECOGNIZES ONE OR
 CC MORE DOMAINS WITHIN THE ALTERNATIVELY SPLICED CS-1 REGION OF
 CC FIBRONECTIN. INTEGRIN ALPHA-4/BETA-7 IS ALSO A RECEPTOR FOR
 CC MADCAM1 AND VCAM1. IT RECOGNIZES THE SEQUENCE L-D-T IN MADCAM1.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. BETA-7
 CC ASSOCIATES WITH EITHER ALPHA-4 OR ALPHA-E.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE INTEGRIN BETA CHAIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 WFMA-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M68903, AAA02749.1,
 DR EMBL: M95632, AAA39323.1,
 DR EMBL: M95633, AAA39324.1,
 DR EMBL: S44607, AAB23193.1,
 DR EMBL: S44604, AAB27396.1,
 DR PIR: P00017, P00017,
 DR PIR: A46271, A46271,
 DR PIR: A42483, A42483,
 DR HSSP: P05106, IJV2,
 DR MCD: MGI:96616, Itgb7,
 DR InterPro: IPR000561, EGF-like,
 DR InterPro: IPR002369, Integrin_B,
 DR InterPro: IPR001169, Integrin_Beta_7,
 DR InterPro: IPR003659, Plexin-like,
 DR InterPro: IPR002035, WFMA-like,
 DR Pfam: PF00362, Integrin_B_1,
 DR PRINTS: PR01186, INTEGRIN,
 DR PRODOM: PD001811, Integrin_B_1,
 DR SMART: SM00187, INB_1,
 DR SMART: SM00423, PSI_1,
 DR SMART: SM00327, VWA_1,
 DR PROSITE: PS00243, INTEGRIN_BETA_3,
 DR PROSITE: PS00022, EGF_1, UNKNOWN_4,
 DR PROSITE: PS01166, EGF_2, UNKNOWN_1,
 KM Integrin: Cell adhesion; Receptor; Transmembrane; Glycoprotein;
 KM Repeat: Signal; Phosphorylation.
 FT SIGNAL 1 19
 FT CHAIN 20 806 INTEGRIN BETA-7,
 FT TRANSSEM 725 745 EXTRACELLULAR (POTENTIAL),
 FT TRANSSEM 746 806 POTENTIAL,
 FT TRANSSEM 150 389 CYTOPLASMIC (POTENTIAL),
 FT TRANSSEM 478 640 WFMA-LIKE,
 FT TRANSSEM 478 526 4 CYSTEINE-RICH TANDEM REPEATS,
 FT TRANSSEM 527 565 1,
 FT TRANSSEM 566 604 11,
 FT TRANSSEM 605 640 1V.
 FT REPEAT 605 640

FT DISULFID 45 476 BY SIMILARITY.
 FT DISULFID 51 61 BY SIMILARITY.
 FT DISULFID 54 91 BY SIMILARITY.
 FT DISULFID 64 80 BY SIMILARITY.
 FT DISULFID 216 223 BY SIMILARITY.
 FT DISULFID 271 311 BY SIMILARITY.
 FT DISULFID 412 428 BY SIMILARITY.
 FT DISULFID 448 688 BY SIMILARITY.
 FT DISULFID 474 478 BY SIMILARITY.
 FT DISULFID 488 500 BY SIMILARITY.
 FT DISULFID 497 537 BY SIMILARITY.
 FT DISULFID 502 511 BY SIMILARITY.
 FT DISULFID 513 527 BY SIMILARITY.
 FT DISULFID 543 548 BY SIMILARITY.
 FT DISULFID 545 574 BY SIMILARITY.
 FT DISULFID 550 559 BY SIMILARITY.
 FT DISULFID 561 566 BY SIMILARITY.
 FT DISULFID 580 585 BY SIMILARITY.
 FT DISULFID 582 613 BY SIMILARITY.
 FT DISULFID 587 596 BY SIMILARITY.
 FT DISULFID 598 605 BY SIMILARITY.
 FT DISULFID 619 624 BY SIMILARITY.
 FT DISULFID 621 666 BY SIMILARITY.
 FT DISULFID 626 635 BY SIMILARITY.
 FT DISULFID 638 641 BY SIMILARITY.
 FT DISULFID 645 654 BY SIMILARITY.
 FT MOD_RES 777 777 PHOSPHORYLATION (BY SIMILARITY).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC: ...) (P0115).
 FT CARBOHYD 250 250 N-LINKED (GLCNAC: ...) (P0115).
 FT CARBOHYD 279 279 N-LINKED (GLCNAC: ...) (P0115).
 FT CARBOHYD 434 434 N-LINKED (GLCNAC: ...) (P0115).
 FT CARBOHYD 531 531 N-LINKED (GLCNAC: ...) (P0115).
 FT CARBOHYD 590 590 N-LINKED (GLCNAC: ...) (P0115).
 FT CARBOHYD 665 665 N-LINKED (GLCNAC: ...) (P0115).
 FT CARBOHYD 674 674 N-LINKED (GLCNAC: ...) (P0115).
 FT CONFLICT 81 81 A -> E (IN REF. 4).
 FT CONFLICT 81 81 A -> G (IN REF. 3).
 FT CONFLICT 124 124 MISSING (IN REF. 2).
 FT CONFLICT 538 538 S -> H (IN REF. 4).
 FT CONFLICT 557 557 R -> H (IN REF. 2).
 SQ SEQUENCE 806 AA, 87411 MW, 173545 GRAVITY 18.6 (GRAV).
 Query Match 2.28; Score 7; DB 1; Length 609
 Best Local Similarity 100.00; Pred. No. 69;
 Matches 7; Conservative 0; Mismatches 0; Indels 0
 OY 71 AVEGLSE 77
 D6 367 AVEGLSE 373
 RESULT 33
 ID GLK1_DROME STANDARD: PFI: 941 AA.
 AC 003445;
 DT 15-JUL-1998 (Ref. 36, Created)
 DT 15-JUL-1998 (Ref. 36, Last sequence update)
 DT 15-JUN-2002 (Ref. 41, Last annotation update)
 DE Glutamate receptor 1 precursor (Kainate-soluble
 DE glutamate receptor).
 GN GLO-R1.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; 66
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-9306263; PubMed-1359540;
 RA Ullrich A., Schuster C.M., Janse P., Schmitt
 RT Kainate-selective subunit expressed in the central nervous
 RL Proc. Natl. Acad. Sci. U.S.A. 89:10484-10488(1992).

CC -1- FUNCTION: L-Glutamate acts as an excitatory neurotransmitter at
 CC MANY SYNAPSES IN THE CENTRAL NERVOUS SYSTEM. THE POSTSYNAPTIC
 CC ACTIONS OF GLU ARE MEDIATED BY A VARIETY OF RECEPTORS THAT ARE
 CC NAMED ACCORDING TO THEIR SELECTIVE AGONISTS (BY SIMILARITY).
 CC -1- SUBUNIT: FORMS HOMOPOLYMERIC CHANNELS WHICH ARE ACTIVATED BY
 CC KAINATE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: CENTRAL NERVOUS SYSTEM.
 CC -1- DEVELOPMENTAL STAGE: NO EXPRESSION IS SEEN IN EARLY EMBRYOGENESIS.
 CC WHEREAS HIGH EXPRESSION OCCURS IN LATE EMBRYOS DURING LARVAL
 CC DEVELOPMENT, EXPRESSION DECREASES TO UNDETECTABLE LEVELS IN LATE
 CC LARVAE, RESUMES AT THE EARLY PUPAL STAGE AND GRADUALLY INCREASES
 CC IN LATE PUPAE AND EARLY ADULT FLIES. HIGH LEVELS OF EXPRESSION
 CC COINCIDE WITH MAJOR STAGES OF NEUROGENESIS.
 CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by not-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.slb.ch/submit/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 CC EMBL: M97192; AAC8575.1; ..
 CC HSSP: P14491; 1GR2.
 CC FLYBAC: FBga0004619; Glu-R1.
 CC InterPro: IPR001828; ANF_receptor.
 CC InterPro: IPR001320; Ion_glu_receptor.
 CC InterPro: IPR001311; SBP/41m_receptor.
 CC Pfam: PF00060; 11q_chan. 1.
 CC Pfam: PF01094; ANF_receptor. 1.
 CC ProDom: PD000500; Ion_glu_receptor. 1.
 CC SMART: SM00079; PBPp. 1.
 CC Receptor: Postsynaptic membrane; Ionic channel; Transmembrane;
 KW Signal: Glycoprotein.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 991 GLUTAMATE RECEPTOR 1
 FT DOMAIN 28 611 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 612 632 POTENTIAL.
 FT TRANSMEM 711 731 POTENTIAL.
 FT TRANSMEM 896 916 POTENTIAL.
 FT DOMAIN 28 32 POLY-GLN.
 FT DOMAIN 367 371 POLY-SER.
 FT CARBOHYD 67 67 N-LINKED (GLCNAC) (POTENTIAL)
 FT CARBOHYD 145 145 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 208 208 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 281 281 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 376 376 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 385 385 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 426 426 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 437 437 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 477 477 N-LINKED (GLCNAC) (POTENTIAL).
 SQ SEQUENCE 991 AA; 111655 MW; F81D5D6D614D0D2C CRC64;
 Query Match: 2.28 Score 7. DB: 1. Length: 991;
 Best Local Similarity: 100.0%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 174 Q01YQEL 180
 DB 185 Q01YQEL 191
 RESULT 44
 ID POLG-IBVHV STANDARD: PPT: 414 AA.
 AC Q01299;
 DI 01-OCT-1996 (Ref. 34, Crea+ed)
 DI 01-OCT-1996 (Ref. 34, Last sequence update)
 DI 15-JUN-2002 (Ref. 41, Last annotation update)
 DE Genome F-1 protein (Contine). Capsid protein c (Cp) protein. Matrix
 DE protein (Envelope protein M); Major envelope protein F; Nonstructural

DE Proteins NS1, NS2A, NS2B, NS4A and NS4B, Protease/helicase
 DE (E:3.4.21.96) (NS3); RNA-directed RNA polymerase (E:2.7.7.48)
 DE (NS5).
 DE Tick-borne encephalitis virus (strain Hypr) (TBEV).
 CC Viruses, ssRNA positive-strand viruses, to DNA stage, Flaviviridae;
 CC Flavivirus.
 CC NCBI_TaxID:70733;
 CC [1]
 RP SEQUENCE FROM N.A.
 PA Walther G., Mandl C.W., Peyer M., Holzmann H., Stiasny K.,
 PA Kunz C., Heinz F.X.;
 RI Submitted (Oct-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 3457-3414 FROM N.A.
 RX MEDLINE:9103656; PubMed:1712858;
 RA Mandl C.W., Kunz C., Heinz F.X.;
 RT "Presence of poly(A) in a flavivirus: significant differences between
 RT the 3' noncoding regions of the genomic RNAs of tick-borne
 RT encephalitis virus strains";
 RL J. Virol. 69:4070-4077(1991).
 CC -1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
 CC HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
 CC NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polyprotein, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1'.
 CC -1- CATALYTIC ACTIVITY: Nucleoside triphosphate - N diphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
 CC LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
 CC PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
 CC PROTEIN C AND NPNA.
 CC -1- MISCELLANEOUS: THE NONSTRUCTURAL PROTEINS NS1 PRESENTS TWO
 CC ALTERNATIVE CLEAVAGE SITES FOR ITS C-TERMINUS, WHICH MAY DEFINE A
 CC SOLUBLE OR A MEMBRANE-BOUND FORM OF NS1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by not-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.slb.ch/submit/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 CC EMBL: U34292; AAB54095.1; ..
 CC EMBL: M76660; AAA47904.1; ..
 CC HSSP: P14316; 1SVB.
 CC InterPro: IPR001410; DEAD.
 CC InterPro: IPR000066; FLAVI_M.
 CC InterPro: IPR001157; FLAVI_NS1.
 CC InterPro: IPR020752; FLAVI_NS2A.
 CC InterPro: IPR000487; FLAVI_NS2B.
 CC InterPro: IPR000104; FLAVI_NS4A.
 CC InterPro: IPR001528; FLAVI_NS4B.
 CC InterPro: IPR000208; FLAVI_NS5.
 DE InterPro: IPR001122; FLAVI_Capsid.
 DE InterPro: IPR050346; FLAVI_glycoprot.
 DE InterPro: IPR001850; FLAVI_helicase.
 DE InterPro: IPR002545; FLAVI_protop.
 DE InterPro: IPR002877; FtsI.
 DE InterPro: IPR001650; Helicase_C.
 DR Pfam: PF00271; helicase_C. 1.
 DR Pfam: PF00869; Flavi_glycoprot. 1.
 DR Pfam: PF00948; Flavi_NS1. 1.
 DR Pfam: PF00949; Flavi_helicase. 1.
 DR Pfam: PF00972; Flavi_NS5. 1.
 DR Pfam: PF01982; Flavi_NS2B. 1.
 DR Pfam: PF01003; Flavi_Capsid. 1.
 DR Pfam: PF01964; Flavi_M. 1.
 DR Pfam: PF01965; Flavi_NS2A. 1.
 DR Pfam: PF01349; Flavi_NS4B. 1.
 DR Pfam: PF01450; Flavi_NS4A. 1.
 DR Pfam: PF01570; Flavi_protop. 1.

DR Pfam: PF01282; FtsY_1.
 DR Pfam: PF02832; Flavl-glycoep-C_1.
 DR Pfam: PF001496; Flavl-NS1_1.
 DR Pfam: PF001556; Flavl-glycoprotein_1.
 DR SMART: SM00490; Helicase_1.
 KW Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
 KW Core protein; Coat protein; Envelope protein; Hydrolyase; Helicase;
 KW ATP-binding; Transmembrane; Nonstructural protein.
 FT INIT_Met 1
 FT CHAIN 1 112
 FT PROPEP 113 205
 FT CHAIN 206 280
 FT CHAIN 281 776
 FT CHAIN 777 7128
 FT CHAIN 7129 1158
 FT CHAIN 1159 1489
 FT CHAIN 1490 2110
 FT CHAIN 2111 2259
 FT CHAIN 2260 2511
 FT CHAIN 2512 3414
 FT NP_BIND 1688 1695
 FT SITE 1779 1782
 FT TRANSMEM 101 112
 FT TRANSMEM 247 259
 FT TRANSMEM 266 280
 FT TRANSMEM 748 751
 FT DISULFID 283 310
 FT DISULFID 340 396
 FT DISULFID 354 385
 FT DISULFID 372 401
 FT DISULFID 466 570
 FT DISULFID 587 618
 FT CARBOHYD 144 144
 FT CARBOHYD 434 434
 FT CARBOHYD 861 861
 FT CARBOHYD 983 983
 FT CARBOHYD 999 999
 FT CARBOHYD 1649 1649
 FT CARBOHYD 1988 1988
 FT CARBOHYD 2044 2044
 FT CARBOHYD 2447 2447
 FT CARBOHYD 2529 2529
 FT CARBOHYD 2726 2726
 SQ SEQUENCE 3414 AA: 378539 MW: EC0B1A5325A08C19 CRG64.

Query Match
 Best Local Similarity 100.0%; Score 7; DB 1; Length 3414.
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 LIAGVLY 38
 DB 761 LIAGVLY 767

RESULT 45
 ID BCEL_HUMAN STANDARD; PRT; 84 AA.
 AC 060756;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Bce-1 Protein.
 GN BCEL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang Z., Huang G.S.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC

CC This SWISS-PROT entry is copyright. It is produced through
 CC between the Swiss Institute of Bioinformatics and the
 CC the European Bioinformatics Institute. There are no restrictions
 CC use by non-profit institutions as long as its copyright
 CC modified and this statement is not removed. Usage by
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/oln/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL: AF068197; AAC18942.1;
 SQ SEQUENCE 84 AA: 9784 MW: D5F7B874FE12422 CRG64;

Query Match
 Best Local Similarity 100.0%; Score 6; DB 1; Length 84.
 Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 28 LSFML 33
 DB 71 LSFML 76

RESULT 36
 ID RK21_PORPU STANDARD; PRT; 104 AA.
 AC P51209;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Chloroplast 50S ribosomal protein L21.
 GN RPL21.
 OS Porphyra purpurea.
 OC Chloroplast.
 OC Eukaryota; Rhodophyta; Bandiophyceae; Bandiiales; Bandiellaceae;
 OX NCBI_TaxID=2787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Keith M.E., Muirhead J.;
 RT "Complete nucleotide sequence of the porphyra purpurea
 RT genome.";
 RL Plant Mol. Biol. Rep. 13:333-336(1995).
 CC OF PROTEIN L20 (BY SIMILARITY).
 CC -1- SIMILARITY: HELIXES TO THE L21 FAMILY OF RIBOSOMAL
 CC -1- SIMILARITY: HELIXES TO THE L21 FAMILY OF RIBOSOMAL
 CC This SWISS-PROT entry is copyright. It is produced through
 CC between the Swiss Institute of Bioinformatics and the
 CC the European Bioinformatics Institute. There are no restrictions
 CC use by non-profit institutions as long as its copyright
 CC modified and this statement is not removed. Usage by
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/oln/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL: U38804; AAC08095.1;
 DR InterPro: IPR001787; Ribosomal_L21P.
 DR Pfam: PF00829; Ribosomal_L21P_1.
 DR Pfam: PF003604; Ribosomal_L21P_1.
 DR TIGRFAMs: TIGR00061; L21_1.
 DR PROSITE: PS0169; RIBOSOMAL_L21_1.
 KW Ribosomal protein; rRNA-binding; Chloroplast.
 SQ SEQUENCE 104 AA: 11877 MW: 768E10785FC4355 CRG64;

Query Match
 Best Local Similarity 100.0%; Score 6; DB 1; Length 104.
 Matches 6; Conservative 0; Mismatches 0; Indels 0;

QY 132 QELTRL 137
 DB 92 QELTRL 97

RESULT 37
 ID C59A_MOUSE STANDARD; PRT; 123 AA.
 AC C59A_MOUSE
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Chloroplast 50S ribosomal protein L21.
 GN RPL21.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Rodentia; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang Z., Huang G.S.;
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 CC

055186;
 16-OCT-2001 (Rel. 40, last sequence update)
 16-OCT-2001 (Rel. 40, last annotation update)
 CD59A glycoprotein precursor (Membrane attack complex inhibition factor) (MACIF) (MAC-inhibitory protein) (MAC-IP) (Protectin).
 CD59A OR CD59
 Mus musculus (Mouse)
 Pukazola, Melissa; Chaudhri, Granita; Vertebrata, Eutelestomi, Mammalia, Eutheria, Rodentia, Sciurognathi; Muridae, Murinae, Mus
 NCBI:taxid=10090;
 11
 SEQUENCE FROM N.A.
 STRAIN=C57BL/6; TISSUE=Kidney;
 MEDLINE=97180887; PubMed=9029105;
 Powell M B, Marchbank K J, Pushner N K, van den Berg C W,
 Morgan B P;
 "Molecular cloning, chromosomal localization, expression, and functional characterization of the mouse analogue of human CD59.";
 J. Immunol. 158:1692-1702(1997).
 12
 SEQUENCE FROM N.A.
 STRAIN=129/Sv.
 MEDLINE=20422499; PubMed=10965140;
 Holt D S, Powell M B, Pushner N K, Morgan B P;
 "Genomic structure and chromosomal location of the gene encoding mouse CD59.";
 Cytogenet. Cell Genet. 85:264-276(2000).
 13
 SEQUENCE FROM N.A.
 STRAIN=C57BL/6; TISSUE=Medulla oblongata, kidney, and placenta;
 MEDLINE=21085660; PubMed=11217851;
 Kawai J, Shingawa A, Shibata K, Yoshino M, Itoh M, Ishi Y,
 Akawa T, Hara A, Fukunishi Y, Kono H, Adachi J, Fukuda S,
 Alzawa K, Izawa M, Nishi K, Kiyosawa H, Kondo S, Yamanka I,
 Saito T, Okazaki Y, Gotohori T, Iwano H, Kasukawa T, Saito F,
 Kadota K, Matsuda H A, Ashburner M, Batalov S, Casavant T,
 Fleischmann W, Gaasterland T, Gissi C, King B, Kochiya H,
 Kiehl J, Lewis S, Matsuo Y, Nikaido I, Pesole G, Quackenbush J,
 Schmitt L M, Stambli F, Suzuki P, Tomita M, Wagner L, Washio T,
 Sakai K, Okido T, Furuno M, Aono H, Balderelli P, Burch G,
 Blake J, Hoffeli D, Hojunga N, Carninci P, de Hernaldo M P,
 Brownstein M J, Bult C, Fierchter C, Fujita M, Gariboldi M,
 Gustavich S, Hill D, Hornann M, Hume D A, Kamiya M, Lee N H,
 Lyons P, Marchionni L, Mashima J, Mazzarelli J, Mombauris P,
 Nordens P, Ping P, Ringwald M, Rodriguez I, Sakamoto N,
 Suzuki H, Sato K, Schoenbach C, Seya T, Shibata Y, Slurch K F,
 Suzuki H, Toyooka K, Wang K H, Welty C, Whitaker C, Wilming L,
 Wynshaw-Boris A, Yoshida K, Hasegawa Y, Kawaji H, Kohlsaki S,
 Hayashizaki Y;
 "Functional annotation of a full-length mouse cDNA collection.";
 Nature 409:685-690(2001).
 14
 FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE ATTACK
 COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE C5 AND/OR C6
 COMPLEXES OF THE ASSEMBLING MAC, THEREBY PREVENTING
 INCORPORATION OF THE MULTIPLE COPIES OF C9 REQUIRED FOR COMPLETE
 FORMATION OF THE OSMOTIC PORE (BY SIMILARITY)
 15
 SUBCELLULAR LOCATION: Attached to the membrane by a GPI anchor.
 16
 TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED (LIVER,
 KIDNEY, SPLEEN, THYMUS, BRAIN AND HEART). LOW LEVELS IN THYMUS
 ALSO EXPRESSED IN MONONUCLEAR CELLS, ERYTHROCYTES AND PLATELETS.
 17
 RARELY DETECTED IN NEUTROPHILS.
 18
 SIMILARITY: CONTAINS 1 UPART/6 DOMAIN
 19
 This SWISS-PROT entry is copyright © 1998 and is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.slb.ch/announcements/>
 or send an email to license@isb.slb.ch).
 DR EMBL: U60473; AAC00055.1;

DR EMBL: AF247652; AAG15414.1;
 DR EMBL: AK018136; BAB31088.1;
 DR EMBL: AF062743; BAB23521.1;
 DR EMBL: AK005507; BAB24087.1;
 DR HSSP: P1987; 1ERG.
 DR MCD: MGI:109177; CD59a.
 DR InterPro: IPR001525; IY6_UPAR.
 DR InterPro: IPR003632; IY-6_CD59.
 DR Pfam: PF00231; UPAR_LY6.1.
 DR ProDom: ProDom129; LY-6_CD59; 1
 DR SMART: SM00134; LY6_1.
 DR PROSITE: PS00983; LY6_UPAR; FALSE_NEG.
 DR Antigen: glycoprotein; GPI-anchor; signal.
 FT SIGNAL 1 23
 FT CHAIN 24 96 CD59A GLYCOPROTEIN.
 FT PROPEP 97 124 PROPEP IN MATURE FORM (BY SIMILARITY).
 FT DOMAIN 24 96 UPAR_LY6.
 FT DISULFID 25 52 BY SIMILARITY.
 FT DISULFID 29 37 BY SIMILARITY.
 FT DISULFID 44 63 BY SIMILARITY.
 FT DISULFID 69 87 BY SIMILARITY.
 FT DISULFID 88 93 BY SIMILARITY.
 FT CARBOHYD 40 40 N-LINKED (GLCNAc...) (POTENTIAL).
 FT CARBOHYD 94 94 N-LINKED (GLCNAc...) (POTENTIAL).
 FT LIPID 95 96 GPI-ANCHOR (BY SIMILARITY).
 SQ SEQUENCE 124 AA; 1644 MW; AABW279P2A7374 CPG64;
 Query Match 1.89, Score 6, DB 1, Length 123;
 Best Local Similarity 100.0%, Freqd. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 36 VLVAAL 41
 111111
 DB 110 VLVAAL 115
 1621-SYNY3 STANFORD; PPT 124 AA.
 AC P74266;
 DT 21-NOV-1997 (Rel. 35, last sequence update)
 DT 01-NOV-1997 (Rel. 35, last annotation update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DB 508 ribosomal protein L21.
 DB PPT OF P74266 SLR1578.
 CS Synchocystis sp. (strain JEC 6803).
 CS Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
 UX NCBI:taxid=1148;
 RN 111
 RP SEQUENCE FROM N.A.
 MEDLINE=97051201; PubMed=8905231;
 RA Kuraku T, Sato S, Kojima H, Tanaka A, Asamizu E, Nakamura Y,
 RA Miyajima N, Hirosewa M, Sugita M, Sasamoto S, Kimura T,
 RA Okumura S, Matsuno A, Moraki A, Nakazaki N, Naruo K,
 RA Okumura S, Shimo S, Takeuchi C, Wada T, Matsumoto A,
 RA Yamada M, Yasuda M, Tabata S;
 "Sequence analysis of the genome of the unicellular cyanobacterium
 Synchocystis sp. strain JEC6803. II. Sequence determination of the
 full-length genome and assignment of potential protein-coding regions.";
 J. Mol. Biol. 310:9-146(1996).
 19
 FUNCTION: THIS PROTEIN BINDS TO 23S RIBOSOMAL RNA IN THE PRESENCE
 OF PROTEIN L20 (BY SIMILARITY).
 20
 SIMILARITY: BELONGS TO THE L21P FAMILY OF RIBOSOMAL PROTEINS.
 21
 This SWISS-PROT entry is copyright © 1998 and is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb.slb.ch/announcements/>
 or send an email to license@isb.slb.ch).
 DR EMBL: D90913; BAB18360.1;

```

DR InterPro: IPR001787; Ribosomal_L21p.
DR Pfam: PF00829; Ribosomal_L21p.1.
DR ProDom: PD003604; Ribosomal_L21p.1.
DR TrEMBL: TIGR00061; L21.1.
DR PROSITE: PS01169; RIBOSOMAL_L21.1.
DR Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 124 AA; 13669 MM; 0201A7C57AFEB0D CRC64;

Query Match 1.8%; Score 6; DB 1; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.3e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0 Gaps 0;

QY 109 QELIRL 114
DB 92 QELIRL 97

RESULT 49
AAA_TRIAH STANDARD; PRT; 125 AA.
AC P81113.
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alboagregin A subunit 3.
OS Trimeresurus albolabris (White-lipped pit viper),
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; ScleroGLOSSA; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID:8765;
RN [1]
RP TISSUE:Venom;
RC MEDLINE=98189535; PubMed=9531050;
RA Kowalska M.A., Tan L., Holt J.C., Peng M., Karczewski J.,
RT Calvege J.J., Niewiarowski S.;
RT *Alboagregins A and B: Structure and interaction with human
RT platelets.*;
RL Thromb Haemost 79:609-613(1998)
CC -!- FUNCTION: BINDS TO PLATELET GPIIb/IIIa RECEPTOR SYSTEM AND STIMULATES
CC AGGREGATION.
CC -!- SUBUNIT: HETEROTRIMER OF THE SUBUNITS 1, 2, 3 AND 4,
CC DISULFIDE-LINKED.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN
DR HSSP: P23807.11XX.
DR InterPro: IPR001304; LECTIN_C.
DR Pfam: PF00059; LECTIN_C.1.
DR SMART: SM00034; CLECT.1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; FALSE_NEG.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
KM Venom; LECTIN.
FT DOMAIN 3 123 C-TYPE LECTIN (LONG FORM).
FT DISULFID 4 15 BY SIMILARITY.
FT DISULFID 32 121 BY SIMILARITY.
FT DISULFID 98 113 BY SIMILARITY.
SQ SEQUENCE 125 AA; 14798 MM; CAFE24C09BDF3293 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 125;
Best Local Similarity 100.0%; Pred. No. 1.3e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 WNGLSL 258
DB 67 WNGLSL 72

RESULT 40
C59B_MOUSE STANDARD; PRT; 129 AA.
AC P58019.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CD59b glycoprotein precursor (Membrane attack complex in-
DE factor) (MACIF) (MAC-inhibitory protein) (MAC-IP) (Pro-
DE CD59b).
GN CD59B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute-
OC leostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus
OC [NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv.
RA MEDLINE=20405635; PubMed 10946279;
RA Qian Y.-M., Qin X., Miwa T., Sun X., Halperin J.A., Song
RA *Identification and functional characterization of a nov-
RT the mouse terminal complement inhibitor CD59.*;
RL J. Immunol. 165:2528-2534(2000).
CC -!- FUNCTION: POTENT INHIBITOR OF THE COMPLEMENT MEMBRANE
CC COMPLEX (MAC) ACTION. ACTS BY BINDING TO THE CR ANTIGEN/
CC COMPLEMENTS OF THE ASSEMBLING MAC, THEREBY PREVENTING
CC INCORPORATION OF THE MULTIPLE COPIES OF C9 REQUIRED FOR
CC FORMATION OF THE OSMOTIC PORE (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a
CC -!- TISSUE SPECIFICITY: SPECIFICALLY IN TESTIS.
DR HSSP: P13687.1CDS.
DR MGD: MG1:1888966; Cd59b.
DR InterPro: IPR001526; LY6_UPAR.
DR InterPro: IPR003632; LY-6_CD59.
DR Pfam: PF00021; UPAR_LY6.1.
DR ProDom: PD003128; LY-6_CD59.1.
DR SMART: SM00134; LY:1.
DR PROSITE: PS00483; LY6_UPAR; FALSE_NEG.
KM Antigen; Glycoprotein; GPI-anchor; Signal.
FT SIGNAL 1 23 PEPTIDAL.
FT CHAIN 24 107 C59B GLYCOPROTEIN.
FT PROPEP 108 129 REMOVED IN MATURE FORM (BY
FT DISULFID 24 107 UPAR/LY6.
FT DISULFID 26 49 BY SIMILARITY.
FT DISULFID 23 36 BY SIMILARITY.
FT DISULFID 42 62 BY SIMILARITY.
FT DISULFID 68 86 BY SIMILARITY.
FT DISULFID 87 92 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC...) (PRO-
FT LIPID 107 107 GPI-ANCHOR (BY SIMILARITY);
SQ SEQUENCE 129 AA; 14266 MM; 2561906A9A9C9CE CRC64;

Query Match 1.8%; Score 6; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.3e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 VLVAIL 41
DB 120 VLVAIL 125

RESULT 41
CYCP_RHOGE STANDARD; PRT; 129 AA.
AC P00142.
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Cytochrome c.
OS Rhodocytus gelatinosus (Rhodospirillum rubrum).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC Roubivivax.
OX NCBI_TaxID=28068;
RN [1]
RP SEQUENCE.
RC MEDLINE=79199608; PubMed=221823;
RA Ambler R.P., Meyer T.E., Kamen M.D.;
RT *Anomalies in amino acid sequences of small cytochromes
RT cytochromes c' from two species of purple phototrophic

```

```

R1  Nature 278:661-662(1979).
R2  [12]
R3  STRUCTURE BY NMR.
R4  MEDLINE 9316176; PubMed8380709;
R5  Portol 1, Cort G, Luchinat C, Villa A J.;
R6  "One- and two-dimensional NMR characterization of oxidized and
R7  reduced cytochrome c' from Rhodocyclus gelatinosus.";
R8  Biochemistry 32:776-783(1993).
R9  [3]
R10 X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS)
R11 Archer M., Banci L., Dikaya E., Poma M.;
R12 "Crystal structure of cytochrome c' from Rhodocyclus gelatinosus and
R13 comparison with other cytochromes c'.";
R14 J. Biol. Inorg. Chem. 2:611-622(1997).
R15 -1- FUNCTION: CYTOCHROME C' IS THE MOST WIDELY OCCURRING BACTERIAL
R16 C-TYPE CYTOCHROME. CYTOCHROMES C' ARE HIGH-SPIN PROTEINS AND THE
R17 HEME HAS NO SIXTH LIGAND. THEIR EXACT FUNCTION IS NOT KNOWN.
R18 PIR: A00135; CCRFCG.
R19 PUB: 1JAE; 21-JAN-98.
R20 InterPro: IPR000345; CytC_heme_bind.
R21 InterPro: IPR002321; Cyt_C11.
R22 Pfam: PF01322; Cytochrome_C_2; 1.
R23 PRINTS: PR00608; CYTOCHROME_C11.
R24 PRODOM: PM004828; Cyt_C11; 1.
R25 PROSITE: PS00190; CYTOCHROME_C; 1.
R26 Electron transport; Heme; 3D-structure.
R27 BINDING 119 119 HEME (COVALENT).
R28 BINDING 122 122 HEME (COVALENT).
R29 METAL 123 123 IRON (HEME AXIAL LIGAND).
R30 SEQUENCE 129 AA; 1528 MW; 05F524E108BA77 CRC64.

Query Match 1.8%; Score 6; DB 1; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 115 KAANGE 120
    111111
DB 109 KAANGE 114

RESULT 42
ID KLA7A_P1G STANDARD: PRT 132 AA
AC Q93776;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 60S ribosomal protein L7a (Surfeit locus protein 3) (Fragment).
GN PRL7A OR SUPP3.
OS Sus scrofa (Pig).
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Mammalia,
OC Mammalia, Eutheria, Cetartiodactyla, Suina, Suidae, Sus.
OC NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RV MEDLINE 94327607; PubMed-8672129;
RA Winterer A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
RT library: analysis of 839 clones.";
RL Mamm. Genome 7:509-517(1996).
R15 -1- SIMILARITY: BELONGS TO THE L7AE FAMILY OF RIBOSOMAL PROTEINS.
R16 -----
R17 This SWISS-PROT entry is copyright. It is produced through a collaboration
R18 between the Swiss Institute of Bioinformatics and the EMBL outstation -
R19 the European Bioinformatics Institute. There are no restrictions on its
R20 use by non-profit institutions as long as its content is in no way
R21 modified and this statement is not removed. Usage by and for commercial
R22 entities requires a license agreement (See http://www.isb-sib.ch)
R23 or send an email to license@isb-sib.ch.
R24 -----
R25 PMID: 814581; CAA2136.1;
R26 InterPro: IPR004038; Ribosomal_L7A

```

```

DR InterPro: IPR004037; Ribosomal_L7Ae.
DR Pfam: PF01248; Ribosomal_L7Ae; 1.
DR PROSITE: PS01082; RIBOSOMAL_L7AE; 1.
KW Ribosomal protein.
FT NON_TER 1 1
FT NON_TER 132 132
SU SEQUENCE 140 AA; 14571 MW; A0010BA44E350B8 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 AGAVI 232
    111111
DB 63 AGAVI 68

RESULT 43
ID ZG49_XENIA STANDARD: PRT 140 AA.
AC P18724;
DT 01 NOV 1990 (Rel. 16, Created)
DT 01 NOV 1990 (Rel. 16, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Castula zinc finger protein XLOC649.1 (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Mammalia,
OC Amphibia, Batrachia, Anura, Mesobatrachia, Pipiloidea, Pipidae.
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RV MEDLINE 940028; PubMed 259712;
RA Neefield W., El-Baradi T., Montzel H., Pieler T., Konster M.,
RA Poeling A., Knoechel W.;
RT "Second-order repeats in Xenopus laevis finger proteins.";
RL J. Mol. Biol. 208:639-659(1989).
DR PIR: S06574; S06574.
DR HSSP: P08046; 1A1H.
DR InterPro: IPR000822; Znf_C2H2.
DR Pfam: PR00946; Zf_C2H2; 5.
DR SMART: SM00455; Znf_C2H2; 5.
DE PROSITE: PS00028; ZINC_FINGER_C2H2_1; 5.
DE PROSITE: PS0157; ZINC_FINGER_C2H2_2; 5.
KW Zinc-finger; Metal-binding; DNA-binding; Repeat.
FT NON_TER 1 1
FT ZN_FING 6 32 C2H2-TYPE.
FT ZN_FING 34 56 C2H2-TYPE.
FT ZN_FING 62 84 C2H2-TYPE.
FT ZN_FING 90 112 C2H2-TYPE.
FT ZN_FING 116 140 C2H2-TYPE.
FT NON_TER 140 140
SU SEQUENCE 140 AA; 16386 MW; EFA270A81583E277 CRC64;

Query Match 1.8%; Score 6; DB 1; Length 140;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 PERKL 104
    111111
DB 100 PERKL 105

RESULT 44
ID H8AL_TORMA STANDARD: PRT 141 AA.
AC T02344;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15 JUN 2002 (Rel. 41, Last annotation update)
DE Hemoglobin alpha-1 chain.
OS Torpedo marmorata (Marbled electric ray).
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Chondrichthyes;

```

CC Elasmobranchii: Squala, Hypnosquala, Pristiogad, Batoidae;
 CC Torpediniformes: Torpedinoidae; Torpedinidae; Torpedo.
 CC NCBI_TaxID=7788;
 CC [1]
 CC SEQUENCE:
 CC MEDLINE=90074179; PubMed=2500465;
 CC Huber F., Braunltzer G.,
 CC "The primary structure of electric ray hemoglobin (Torpedo marmorata).
 CC Bohr effect and phosphate interaction.";
 CC Biol. Chem. Hoppe-Seyler 370:831-838(1989)
 CC -1- FUNCTION: Involved in oxygen transport from the lung to the
 CC various peripheral tissues.
 CC -1- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.
 CC -1- TISSUE SPECIFICITY: Red blood cells.
 CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.
 CC PIR: S05418; HARRY.
 CC HSSP: p56691; ICG5.
 CC InterPro: IPR000971; Globin.
 CC Pfam: PF00042; globin; 1.
 CC PROSITE: PS01033; GLOBIN; 1.
 CC Heme: oxygen transport; Transport; Erythrocyte.
 CC METAL: 59 59 IRON (HEME DISTAL LIGAND).
 CC METAL: 88 88 IRON (HEME PROXIMAL LIGAND).
 CC SEQUENCE 141 AA; 15866 MW; 31BA7A1756FC06B7 CPO64;

Query Match 1.88; Score 6; DR 1; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 HGALVL 25
 111111
 DB 59 HGALVL 64

RESULT 45
 HBF1_URECA STANDARD; PRT; 141 AA.
 ID HBF1_URECA
 AC P06148;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hemoglobin F-1.
 OS Urechis caupo (Inkeeper worm) (Spoonworm).
 CC Eukaryota, Metazoa, Echinoda, Xenopneustae; Urechidae; Urechis.
 CC NCBI_TaxID=6431;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=87057334; PubMed=3782128;
 CC Garey J.R., Riggs A.F.;
 CC "The hemoglobin of Urechis caupo. The cDNA-derived amino acid
 CC sequence.";
 CC J. Biol. Chem. 261:16446-16450(1986).
 CC [2]
 CC X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 CC MEDLINE=92384990; PubMed=1515107;
 CC Kolatkar P.R., Ernst S.R., Hackert M.L., Ogata C.M.,
 CC Hendrickson W.A., Merril E.A., Philackerley R.P.;
 CC "Structure determination and refinement of homotetrameric hemoglobin
 CC from Urechis caupo at 2.5-A resolution.";
 CC Acta Crystallogr. B 48:191-199(1992).
 CC [3]
 CC X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 CC MEDLINE=94180394; PubMed=8133523;
 CC Kolatkar P.R., Hackert M.L., Riggs A.F.;
 CC "Structural analysis of Urechis caupo hemoglobin.";
 CC J. Mol. Biol. 237:87-97(1994).
 CC -1- FUNCTION: HEMOGLOBIN F-1 APPEARS TO FUNCTION IN STORAGE, FATHER
 CC THAN TRANSPORT OF OXYGEN.
 CC -1- SUBUNIT: Homotetramer.
 CC -1- MISCELLANEOUS: HEMOGLOBIN F-1 BINDS OXYGEN NONCOOPERATIVELY AND
 CC ALMOST INDEPENDENTLY OF PH.
 CC -1- SIMILARITY: BELONGS TO THE GLOBIN FAMILY.

CC This SWISS-Prot entry is copyright. It is produced through
 CC the European Bioinformatics Institute. There are no fees
 CC use by non-profit institutions as long as its content
 CC modified and this statement is not removed. Usage by
 CC entities requires a license agreement (see <http://www.ebi.ac.uk>
 CC or send an email to license@ebi.ac.uk).
 CC EMBL: J02624; AAA30331.1;
 CC PIR: A25537; A25537;
 CC PDB: 1UTH; 3I-OCF-93.
 CC InterPro: IPR000971; Globin.
 CC Pfam: PF00042; globin; 1.
 CC PROSITE: PS01033; GLOBIN; 1.
 CC Heme: oxygen transport; Transport; Erythrocyte.
 CC 3D-structure.
 CC FT INIT_MET 0 0 IRON (HEME PROXIMAL LIGAND)
 CC FT METAL 94 94
 CC FT HELIX 4 17
 CC FT TURN 18 18
 CC FT HELIX 19 36
 CC FT HELIX 38 43
 CC FT TURN 45 49
 CC FT HELIX 52 57
 CC FT HELIX 59 77
 CC FT TURN 78 82
 CC FT HELIX 83 88
 CC FT TURN 89 90
 CC FT HELIX 91 95
 CC FT TURN 96 98
 CC FT HELIX 101 118
 CC FT HELIX 123 138
 CC FT TURN 139 140
 CC SQ SEQUENCE 141 AA; 15062 MW; 2F68H507903E646A CPO64;

Query Match 1.88; Score 6; DR 1; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0;

OY 34 AGVLVA 39
 111111
 DB 133 AGVLVA 138

Search completed: December 7, 2002, 11:16:09
 Job time : 28 secs

100

GenCore version 5.1.1
Copyright (c) 1993 - 2002 CompuGen Ltd

OM protein - protein search, using sw model

Run on: December 7, 2002. 11:08:20, Search time: 33 Seconds

(Without alignments)
2029.252 Million cell updates/sec

Title: US-09-831-458A-12

Perfect score: 325
Sequence: 1 MSDSKEDPVQQLGLGCLJH
PQVDNPNYIQQKPAACFPDE 325

Scoring table: 01,100

Searched: 671580 seqs. 206047115 residues

Word size : 5

Total number of hits satisfying chosen parameters: 33422

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

```
Database : SPTREMBL_21:*
```

```

1:  sp.archaei:21:*
2:  sp.bacteria:*
3:  sp.fungi:*
4:  sp.human:*
5:  sp.invertebrate:*
6:  sp.mammal:*
7:  sp.mmc:*
8:  sp.organellae:*
9:  sp.phage:*
10: sp.plant:*
11: sp.podot:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeop:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	168	51.7	296	4	Q9Y2A8	Q9Y2A8 homo sapien
2	165	51.1	376	4	Q9BKS3	Q9BKS3 homo sapien
3	166	51.1	399	4	Q9H2X3	Q9H2X3 homo sapien
4	153	47.1	272	4	Q969M4	Q969M4 homo sapien
5	153	47.1	399	4	Q9H2O9	Q9H2O9 homo sapien
6	138	42.5	263	4	Q962P3	Q962P3 homo sapien
7	138	42.5	332	4	Q960P5	Q960P5 homo sapien
8	125	38.5	251	4	Q960P4	Q960P4 homo sapien
9	120	36.9	232	4	Q960P6	Q960P6 homo sapien
10	97	29.8	216	4	Q9H8F0	Q9H8F0 homo sapien
11	78	24.0	148	4	Q960P7	Q960P7 homo sapien
12	78	24.0	297	4	Q960P8	Q960P8 homo sapien
13	78	24.0	312	4	Q960O7	Q960O7 homo sapien
14	78	24.0	334	4	Q960P9	Q960P9 homo sapien
15	78	24.0	360	4	Q960O4	Q960O4 homo sapien
16	78	24.0	380	4	Q960O5	Q960O5 homo sapien

[illegible]

APPENDIX

RESULT 1

ID	Q9Y2A
----	-------

PRELIMINARY;

PORT; 256 AA.

DT 01-NOV-1999 (TREMBLrel. 12, created)
 DT 01-NOV-1999 (TREMBLrel. 12, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Type II membrane protein similar to HIV gp120-binding site
 DE lectin.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo
 OX NCBI_Taxid:9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MPRLIN=99171880; PubMed:19072769;
 RA Yokoyama-Kobayashi M., Sakane S., Kato S.,
 RT "Selection of cDNAs encoding putative type II membrane p
 RL cell surface from a human full-length cDNA bank".
 RL Gene 228:161-167(1999).
 DR EMBL; AB015629; BAA76496.1;
 DR InterPro: IPR001304; lectin_C;
 DR Pfam: PFD0054; lectin_C; 1
 SJ SEQUENCE 299 AA; 33527 MW; 762285815E20E54F C7C7C4;

```

DB 123 LPEKSLCELYOELTFLKAAVAFILPPEKSLCELYOELTFLKAAVAFILP 169
|||||
RESULT 2
Q9BXK3 PRELIMINARY: PRT: 376 AA.
AC Q9BXK3:
DI 01-JUN-2001 (TREMblrel: 17, created)
DI 01-JUN-2001 (TREMblrel: 21, last sequence update)
DI 01-JUN-2002 (TREMblrel: 21, last annotation update)
DE L-SIGN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:21157496; PubMed 11257134;
RA Bashirova A.A., Gelfandov T.R., van Duinbeven G.C.F.,
RA van Vliet S.J., Ellering J.B.G., Martin M.P., Wu L., Martin T.H.,
RA Viebig N., Knolle P.A., Kewalramani V.N., van Kooyk Y., Carrington M.;
KI "A dendritic cell-specific intercellular adhesion molecule 3-grabbing
KI integrin (DC-SIGN)-related protein is highly expressed on human
KI liver sinusoidal endothelial cells and promotes HIV-1 infection."
RL J Exp Med 193:671-678(2001)
DR EMBL: AF290887; AAK20998.1; -.
DR HSP: P22897; IEG3.
DR InterPro: IPR002353; AntiFcecell.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1
DR PRINTS: PR00356; ANTI_FCEZELL.
DR SMART: SM00034; CLECT_1.
DR PROSITE: PS00615; G_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50411; G_TYPE_LECTIN_2; 1
SQ SEQUENCE 376 AA: 42724 MW: 36319991E33BE90 CRC64;

Query Match 51.1%; Score 166; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 2,7e-153;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 160 LKAAGVETLPGSKQGGYVGLTFLKTAPEFLCFRHPKMTFFPGGNYPMNSGPNMHVSV 219
DB 211 LKAAGVETLPGSKQGGYVGLTFLKTAPEFLCFRHPKMTFFPGGNYPMNSGPNMHVSV 270
QY 220 TACQVFAQIVVVKTAEEENFQIQTSPSNRPSWWSGLSDLNQEGJTWQVWDSPLSPFQR 279
DB 271 TACQVFAQIVVVKTAEEENFQIQTSPSNRPSWWSGLSDLNQEGJTWQVWDSPLSPFQR 330
QY 280 YWNSGEFNNSGNEDEAFESGSGWNRKLDVNTWICKKPAACFDE 325
DB 331 YWNSGEFNNSGNEDEAFESGSGWNRKLDVNTWICKKPAACFDE 376

RESULT 3
Q9HXZ3 PRELIMINARY: PRT: 399 AA.
AC Q9HXZ3:
DI 01-MAR-2001 (TREMblrel: 16, created)
DI 01-MAR-2001 (TREMblrel: 16, last sequence update)
DI 01-JUN-2002 (TREMblrel: 21, last annotation update)
DE Probable mannose binding C-type lectin DC-SIGN2 type I
DE isoform.
GN ABC OR CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:23432257; PubMed 10757799;
RA Solliou E.J., Barten R., Trowsdale J.;
DR EMBL: AY042245; AAK91860.1; -.
SQ SEQUENCE 399 AA: 42724 MW: 36319991E33BE90 CRC64;

```

```

RT 19913.7;
RT 1 Immortal 165-2042(2000);
RN [2]
RP SEQUENCE FROM N.A.
RA Solliou E.J.;
RA Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed 11377487;
RA Munnadi S., Catano G., Lam L., Heefle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT specific ICAM-3-grabbing Integrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
RL J Biol Chem 276:33396-33212(2001);
DR EMBL: AF209491; AAC13815.2; -.
DR EMBL: AF209490; AAC13815.2; J01NED.
DR EMBL: AY042244; AAK91859.1; -.
DR HSP: P22897; IEG3.
DR InterPro: IPR002353; AntiFcecell.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1
DR PRINTS: PR00356; ANTI_FCEZELL.
DR SMART: SM00034; CLECT_1.
DR PROSITE: PS00615; G_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50411; G_TYPE_LECTIN_2; 1
DR LectIn.
SQ SEQUENCE 399 AA: 45350 MW: 06A7C9972AFA593 CRC64;

Query Match 51.1%; Score 166; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 2,9e-153;
Matches 166; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 229 TACQVFAQIVVVKTAEEENFQIQTSPSNRPSWWSGLSDLNQEGJTWQVWDSPLSPFQR 279
DB 294 TACQVFAQIVVVKTAPEENFQIQTSPSNRPSWWSGLSDLNQEGJTWQVWDSPLSPFQR 363
QY 280 YWNSGEFNNSGNEDEAFESGSGWNRKLDVNTWICKKPAACFDE 325
DB 354 YWNSGEFNNSGNEDEAFESGSGWNRKLDVNTWICKKPAACFDE 399

RESULT 4
Q969M4 PRELIMINARY: PRT: 272 AA.
AC Q969M4:
DI 01-DEC-2001 (TREMblrel: 19, created)
DI 01-DEC-2001 (TREMblrel: 19, last sequence update)
DI 01-DEC-2001 (TREMblrel: 19, last annotation update)
DE MDC-SIGN2 type IV isoform (MDC-SIGN2 type III isoform).
GN CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed 11377487;
RA Munnadi S., Catano G., Lam L., Heefle A., Telles V., Begum K.,
RA Jimenez F., Ahuja S.S., Ahuja S.K.;
RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT specific ICAM-3-grabbing Integrin 1 (DC-SIGN1) and DC-SIGN2
RT Isoforms Inter-individual Variation in Expression of DC-SIGN
RT Transcripts."
RL J Biol Chem 276:33396-33212(2001);
DR EMBL: AY042245; AAK91860.1; -.
DR EMBL: AY042246; AAK91860.1; -.
SQ SEQUENCE 272 AA: 30481 MW: B3BEF2DE541F5E CRC64;

```

```

Query Match          47.1%; Score 153; DB 4; Length 272;
Best Local Similarity 100.0%; Pred. No. 66-150;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 16 GCLGHALVLDLSPMLAGVLAIVLVQSKVPSLSQESQEDATYQNTLTAKAVGEL 75
      |||||||
Db 44 GCLGHALVLDLSPMLAGVLAIVLVQSKVPSLSQESQEDATYQNTLTAKAVGEL 103
      |||||||

QY 76 SEKSKLOEIVQELTQTKAAVGEPEKSKLOEIVQELTRKAAVGEPEKSKLOEIVQELT 135
      |||||||
Db 104 SEKSKLOEIVQELTQTKAAVGEPEKSKLOEIVQELTRKAAVGEPEKSKLOEIVQELT 163
      |||||||

QY 136 RLKAAVGEPEKSKLOEIVQELTRKAAVGEPE 168
      |||||||
Db 164 RLKAAVGEPEKSKLOEIVQELTRKAAVGEPE 196
      |||||||

RESULT 5
Q9H209 PRELIMINARY; PRT; 399 AA.
AC Q9H209;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable mannose binding C-type lectin DC-SIGNR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:20432267; PubMed:10975799;
RA Solliex E.J.; Batten R.; Tromsdale J.;
RT "DC-SIGN: a related gene, DC-SIGNR" and CD23 form a cluster on
RT 1p13.1;
RL J. Immunol. 165:2937-2942(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Solliex E.J.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF245219; AAG13848.2;
DR HSSP: P22897; IEGG.
DR InterPro: IPR002353; Antifreeze1.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1
DR PRINTS: PR00356; ANTIFREEZE11.
DR SMART: SM00034; CLECT_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
DR KEGG:
KW lectin.
SU SEQUENCE 399 AA; 45357 MW; FRC999F72AE6589 CR664;

Query Match          47.1%; Score 153; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 8-56-150;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 16 GCLGHALVLDLSPMLAGVLAIVLVQSKVPSLSQESQEDATYQNTLTAKAVGEL 75
      |||||||
Db 44 GCLGHALVLDLSPMLAGVLAIVLVQSKVPSLSQESQEDATYQNTLTAKAVGEL 103
      |||||||

QY 76 SEKSKLOEIVQELTQTKAAVGEPEKSKLOEIVQELTRKAAVGEPEKSKLOEIVQELT 145
      |||||||
Db 104 SEKSKLOEIVQELTQTKAAVGEPEKSKLOEIVQELTRKAAVGEPEKSKLOEIVQELT 163
      |||||||

QY 136 RLKAAVGEPEKSKLOEIVQELTRKAAVGEPE 168
      |||||||
Db 164 RLKAAVGEPEKSKLOEIVQELTRKAAVGEPE 196
      |||||||

RESULT 6
Q960P3 PRELIMINARY; PRT; 263 AA.
AC Q960P3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable mannose binding C-type lectin DC-SIGNR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:20432267; PubMed:10975799;
RA Solliex E.J.; Batten R.; Tromsdale J.;
RT "DC-SIGN: a related gene, DC-SIGNR" and CD23 form a cluster on
RT 1p13.1;
RL J. Immunol. 165:2937-2942(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Solliex E.J.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF245219; AAG13848.2;
DR HSSP: P22897; IEGG.
DR InterPro: IPR002353; Antifreeze1.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
DR KEGG:
KW lectin.
SU SEQUENCE 399 AA; 45357 MW; FRC999F72AE6589 CR664;

```

```

DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SDC SIGN2 type III isoform.
GN CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed:11337487;
RA Mummidi S.; Catano G.; Lam I.; Hoelle A.; Telles V.; Best
RA Jimenez F.; Ahuja S.S.; Ahuja S.K.;
RT "Extensive repertoire of Membrane-bound and Soluble Dendritic
RT specific ICAM-3-grabbing Nomenclature 1 (DC-SIGN) and Dendritic
RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
RT Transcripts";
RL J. Biol. Chem. 276:33196-33212(2001).
RL EMBL: AY042440; AAK91865.1;
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1
DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
SU SEQUENCE 263 AA; 30102 MW; 7611576250424CF CR664;

Query Match          42.5%; Score 138; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 26-134;
Matches 138; Conservative 0; Mismatches 0; Indels 0.

QY 188 PFLCPHCPKRWTFYFGNGYFMSNSGPRNHHSVATGCGEVALVYKFAFE 325
      |||||||
Db 126 EKLCPHCPKRWTFYFGNGYFMSNSGPRNHHSVATGCGEVALVYKFAFE 325
      |||||||

QY 248 SNKFSMGLSLDNLQGTQWQWVMSPLSPFORVWNSGPPNNSGNDCAFE 325
      |||||||
Db 186 SNKFSMGLSLDNLQGTQWQWVMSPLSPFORVWNSGPPNNSGNDCAFE 325
      |||||||

QY 308 DVDNWTCKKTPACGRDE 325
      |||||||
Db 246 EVNWTCKKTPACGRDE 246
      |||||||

RESULT 7
Q960P5 PRELIMINARY; PRT; 442 AA.
AC Q960P5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SDC SIGN2 type I isoform.
GN CD209L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed:11337487;
RA Mummidi S.; Catano G.; Lam I.; Hoelle A.; Telles V.; Best
RA Jimenez F.; Ahuja S.S.; Ahuja S.K.;
RT "Extensive repertoire of Membrane-bound and Soluble Dendritic
RT specific ICAM-3-grabbing Nomenclature 1 (DC-SIGN) and Dendritic
RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
RT Transcripts";
RL J. Biol. Chem. 276:33196-33212(2001).
RL EMBL: AY042440; AAK91865.1;
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.1
DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
SU SEQUENCE 312 AA; 37954 MW; 7611576250424CF CR664;

```

Query Match 42.5%; Score 138; DB 4; Length 332;
 Best Local Similarity 100.0%; Pred. No. 2,5e-134;
 Matches 138; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 188 ERLPHPTKFWTPQSNVYFMSNSQPMNHVSIAQVEVPAVLVRIAEKQNFLOQPSR 247
 |||||||
 DB 195 ERLPHPTKFWTPQSNVYFMSNSQPMNHVSIAQVEVPAVLVRIAEKQNFLOQPSR 254
 |||||||
 QY 248 SRFESWMEISDLNPGTGWGKSPISPSFGYKNSGFPNNNGNHVCAEFGSGMNINFG 307
 |||||||
 DB 255 SRFESWMEISDLNPGTGWGKSPISPSFGYKNSGFPNNNGNHVCAEFGSGMNINFG 314
 |||||||
 QY 308 DVDNWKCKKPAACPRDE 325
 |||||||
 DB 315 DVDNWKCKKPAACPRDE 332
 |||||||

RESULT 8

Q960P4 PRELIMINARY; PRT: 251 AA.

AC Q960P4: 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 GN SDC-SIGN2 type II isoform.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA PubMed=1137487;
 RA Mummidi S., Catalan G., Lam L., Hoeftle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms, Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts."
 RL J. Biol. Chem. 276:33196-33202(2001)
 DR EMBL: AY042239; AAK91864.1;
 SO SEQUENCE 251 AA; 28373 MW; 86F69F6CD6055DCE CRC64.

Query Match 38.5%; Score 125; DB 4; Length 251;
 Best Local Similarity 100.0%; Pred. No. 5.8e-171;
 Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 VSKVPSLSQSGEQGATAYNTOLKAAGVPISEKSKLOETIYQELTOLKAAGVPEPSK 103
 |||||||
 DB 51 VSKVPSLSQSGEQGATAYNTOLKAAGVPISEKSKLOETIYQELTOLKAAGVPEPSK 110
 |||||||
 QY 104 LOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAA 163
 |||||||
 DB 111 LOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAA 170
 |||||||
 QY 164 VSELP 168
 |||||
 DB 171 VSELP 175
 |||||

RESULT 9

Q960P6 PRELIMINARY; PRT: 232 AA.

AC Q960P6: 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 GN MDC-SIGN2 type VI isoform.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RA PubMed=1137487;
 RA Mummidi S., Catalan G., Lam L., Hoeftle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms, Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts."
 RL J. Biol. Chem. 276:33196-33202(2001).
 DR EMBL: AY042237; AAK91862.1;
 DR InterPro: IPR001304; IcdIn_C1.
 DR Pfam: PF00059; IcdIn_C1.
 SO SEQUENCE 232 AA; 26255 MW; 19C1B150B2451B72 CRC64;

Query Match 36.9%; Score 120; DB 4; Length 232;
 Best Local Similarity 100.0%; Pred. No. 8.2e-116;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GCLGHCALVQLISPMILAGVIALLVGVSKVPSLSQSGEQGATAYNTOLKAAGVCEL 75
 |||||||
 DB 44 GCLGHCALVQLISPMILAGVIALLVGVSKVPSLSQSGEQGATAYNTOLKAAGVCEL 103
 |||||||
 QY 76 SEKSKLOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAAGVPEPSKLOETIYQELT 135
 |||||||
 DB 104 SEKSKLOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAAGVPEPSKLOETIYQELT 163
 |||||||

RESULT 10

Q9H8F0 PRELIMINARY; PRT: 216 AA.

AC Q9H8F0: 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE CDNA FLV1006118, clone pLACF200002, highly similar to Homo sapiens
 DE type II membrane protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE-PLACENTA;
 RA Ishtar, C., Hagstall, C., Saitama, T., Otsuki, T., Sasaki, T.,
 RA Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H.,
 RA Tanase, T., Nomura, Y., Tojiya, S., Komai, F., Hara, K., Takeuchi, K.,
 RA Arita, M., Nakamura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
 RA Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuno, Y., Oshima, A.;
 RT "NEKO Human cDNA sequencing project."
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK024750; BAB14667.1;
 SO SEQUENCE 216 AA; 24438 MW; 4A550078B8E2C3CB8 CRC64;

Query Match 29.8%; Score 97; DB 4; Length 216;
 Best Local Similarity 100.0%; Pred. No. 5.2e-92;
 Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 VGEISKSNLOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAAGVPEPSKLOETIY 141
 |||||||
 DB 44 VGEISKSNLOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAAGVPEPSKLOETIY 153
 |||||||
 QY 133 VGEISKSNLOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAAGVPEPSKLOETIY 168
 |||||||
 DB 104 VGEISKSNLOETIYQELTOLKAAGVPEPSKLOETIYQELTOLKAAGVPEPSKLOETIY 140
 |||||||

RESULT 11

Q960P7 PRELIMINARY; PRT: 148 AA.

AC Q960P7: 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE SDC-SIGNIB type IV isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=1137487;
 RA Mummidi S., Catano G., Lam L., Hoelle A., Telles V., Begun K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive Repertoire of Membrane-bound and Soluble Pentameric Cell-
 RT Specific ICM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042333; AAK91858.1;
 SO SEQUENCE 148 AA; 16608 MW; 5FDD84AD6DA1AF7 CRC64;

Query Match 24.0%; Score 78; DB 4; Length 148;
 Best Local Similarity 100.0%; Pred No. 1.8e-72;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONLTOLKAAGVGLSKLQEIYQELTOLKAAGVGLPEKSKLQEIYQELTIRLAKA 117
 DB 50 QDAIYONLTOLKAAGVGLSKLQEIYQELTOLKAAGVGLPEKSKLQEIYQELTIRLAKA 109
 QY 118 VGLPEKSKLQEIYQELT 135
 DB 110 VGLPEKSKLQEIYQELT 127

RESULT 12

Q96Q08 PRELIMINARY; PRT; 297 AA.
 ID Q96Q08;
 AC Q96Q08;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE SDC-SIGNIB type III isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=1137487;
 RA Mummidi S., Catano G., Lam L., Hoelle A., Telles V., Begun K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive Repertoire of Membrane-bound and Soluble Pentameric Cell-
 RT Specific ICM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042332; AAK91857.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C; 1;
 SO SEQUENCE 297 AA; 31873 MW; C8C0951B915FC5BF CRC64;

Query Match 24.0%; Score 78; DB 4; Length 297;
 Best Local Similarity 100.0%; Pred No. 3.4e-72;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONLTOLKAAGVGLSKLQEIYQELTOLKAAGVGLPEKSKLQEIYQELTIRLAKA 117
 DB 50 QDAIYONLTOLKAAGVGLSKLQEIYQELTOLKAAGVGLPEKSKLQEIYQELTIRLAKA 109
 QY 118 VGLPEKSKLQEIYQELT 135
 DB 110 VGLPEKSKLQEIYQELT 127

RESULT 13

Q96Q07 PRELIMINARY; PRT; 412 AA.
 ID Q96Q07;
 AC Q96Q07;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE MDC-SIGNIA type III isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=1137487;
 RA Mummidi S., Catano G., Lam L., Hoelle A., Telles V., Begun K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive Repertoire of Membrane-bound and Soluble Pentameric Cell-
 RT Specific ICM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042323; AAK91848.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C; 1;
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00041; C_TYPE_LECTIN_2; 1;
 SO SEQUENCE 312 AA; 35178 MW; 65761618F64ABE CRC64;

Query Match 24.0%; Score 78; DB 4; Length 312;
 Best Local Similarity 100.0%; Pred No. 3.5e-72;
 Matches 78; Conservative 0; Mismatches 0; Indels 0;

QY 58 QDAIYONLTOLKAAGVGLSKLQEIYQELTOLKAAGVGLPEKSKLQEIYQELTIRLAKA 117
 DB 74 QDAIYONLTOLKAAGVGLSKLQEIYQELTOLKAAGVGLPEKSKLQEIYQELTIRLAKA 109
 QY 118 VGLPEKSKLQEIYQELT 135
 DB 134 VGLPEKSKLQEIYQELT 151

RESULT 14

Q96Q09 PRELIMINARY; PRT; 334 AA.
 ID Q96Q09;
 AC Q96Q09;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, last annotation update)
 DE SDC-SIGNIB type II isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=1137487;
 RA Mummidi S., Catano G., Lam L., Hoelle A., Telles V., Begun K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive Repertoire of Membrane-bound and Soluble Pentameric Cell-
 RT Specific ICM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms. Inter-individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042331; AAK91856.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C; 1;
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00041; C_TYPE_LECTIN_2; 1;
 SO SEQUENCE 334 AA; 37843 MW; BE766FC4111CB6BF CRC64;

Query Match	24.0%	Score 78:	PR 4:	Length 380:
Best Local Similarity	100.0%	Prod. No. 4.2e-72:		
Matches 78:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY	58	CGATGCTGTTAAAGGSLPEKSKLQETIYQELTQKAAAGSLPEKSKLQETIYQELTQKAA	117	
RA	111	TT		
RT	50	CGATGCTGTTAAAGGSLPEKSKLQETIYQELTQKAAAGGSLPEKSKLQETIYQELTQKAA	109	
QY	118	VGELPEKSKLQETIYQELT	135	
RA	111	TT		
EB	110	VGELPEKSKLQETIYQELT	127	
RESULT 17				
Q96000		PRELIMINARY:	PR:	380 AA
AC	Q96000:			
DT	01-DEC-2001 (TEMBLrel. 19, (reated)			
DT	01-DEC-2001 (TEMBLrel. 19, last sequence update)			
DT	01 MAR 2002 (TEMBLrel. 20, last annotation update)			
DE	SOC-STGNR type 1 isoform.			
GN	CD209.			
OS	Homo sapiens (Human).			
OC	Eukaryota, Metazoa, Chordata, Cladacea, Vertebrata, Euteleostomi:			
OC	Mammalia, Eutheria, Primates, Catarrhini, Homiidae, Homo.			
OX	NCBI_taxid:9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	PubMed-1137487:			
RA	Mummid S., Catalan G., Lam L., Hoelle A., Telles V., Begum K.,			
RA	Jimenez P., Ahuja S.S., Ahuja S.K. &			
RZ	"Extensive Repertoire of Membrane-bound and Soluble Invariant C-11-			
RT	Specific 16M-3'-trabbing Nucleotide 1 (DC-STGN1) and DC-STGN2			
RT	Isoforms. Inter-individual Variation in Expression of DC-STGN			
PT	Transcripts."			
PL	J. Biol. Chem. 276:33196-33212(2001).			
DR	EMBL: AY042230; AAK91855.1; "			
DR	InterPro: IPR001404; Loclin_C1			
DR	Prfam: PFD00059; Loclin_C1			
DR	PROSITE: PS00615; C_TYPE_LETTIN_1; UNKNOWN_1.			
DR	PROSITE: PS00615; C_TYPE_LETTIN_2; 1.			
SO	SEQUENCE 380 AA; 43340 MW; F3D09F9FH7D044B CRC64;			

AC 096008: 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE MDC-SIGNIA type II isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed:11337487;
 RA Mummidi S., Catano G., Lam L., Hoeftle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive Repertoire of Membrane-Bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042222; AAK91847.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
 SQ SEQUENCE 398 AA: 45031 MW: D9257648679102E CRC64;

Query Match 24.0% Score 78; DB 4; Length 398;
 Best Local Similarity 100.0%; Pred. No. 4, 4e-72;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYQNTLTKAAVGESEKSLQEIYQELTOLKAAGVGLPEKSLQEIYQELTRKAA 117
 DB 74 QDAIYQNTLTKAAVGESEKSLQEIYQELTOLKAAGVGLPEKSLQEIYQELTRKAA 133
 QY 118 VGLPEKSLQEIYQELT 135
 DB 134 VGLPEKSLQEIYQELT 151

RESULT 19

Q96001 PRELIMINARY; PRT; 404 AA.
 AC 096001: 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE MDC-SIGNIB type I isoform.
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed:11337487;
 RA Mummidi S., Catano G., Lam L., Hoeftle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive Repertoire of Membrane-Bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: AY042222; AAK91847.1;
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
 SQ SEQUENCE 404 AA: 45570 MW: E4DB570273D9DC30 CRC64;

Query Match 24.0% Score 78; DB 4; Length 404;
 Best Local Similarity 100.0%; Pred. No. 4, 4e-72;
 Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYQNTLTKAAVGESEKSLQEIYQELTOLKAAGVGLPEKSLQEIYQELTRKAA 117
 DB 74 QDAIYQNTLTKAAVGESEKSLQEIYQELTOLKAAGVGLPEKSLQEIYQELTRKAA 133
 QY 118 VGLPEKSLQEIYQELT 135
 DB 134 VGLPEKSLQEIYQELT 151

RESULT 20

Q9NNX6 PRELIMINARY; PRT; 404 AA.
 AC 09NNX6: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Membrane-associated lectin type-2 (probable mannose-binding
 DE lectin DC-SIGN) (MDC-SIGNIA type I isoform).
 GN CD209.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=PLACENTA;
 RX MEDLINE=92390446; PubMed:1518869;
 RA Curtis B.M., Scharnowski S., Watson A.J.;
 RT *Sequence and expression of a membrane-associated C-type
 RT lectin exhibits C4-independent binding of human immunodeficiency
 RT envelope glycoprotein gp120.*;
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8456-8460(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20432267; PubMed:10975799;
 RA Solileux E.J., Batten R., Townsend J.;
 RT *DC-SIGN, a related gene, DC-SIGNR, and CD23 form a C-type
 RT lectin family.*;
 RL J. Immunol. 165:2937-2942(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21157496; PubMed:1125714;
 RA Bashilova A.A., Gelfand-von H., van Euijnhoven G.C.,
 RA van Vliet S.J., Ellering J.B.G., Martin M.P., Ma L., Ma
 RA Vleider N., Knolle P.A., KewalRamani V.N., van Kooyk Y.,
 RT *A dendritic cell-specific intercellular adhesion molecule-
 RT nonintegrin (DC-SIGN)-related protein is highly expressed
 RT in liver sinusoidal endothelial cells and promotes HIV-1
 RT J. Exp. Med. 193:671-678(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX PubMed:11337487;
 RA Mummidi S., Catano G., Lam L., Hoeftle A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT *Extensive Repertoire of Membrane-Bound and Soluble Dendritic Cell-
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
 RT Isoforms Inter-Individual Variation in Expression of DC-SIGN
 RT Transcripts.*;
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR EMBL: M98457; AAF77072.1;
 DR EMBL: AF209479; AAG13814.1;
 DR EMBL: AF209486; AAK20997.1;
 DR EMBL: AY042221; AAK91846.1;
 DR HSSP: P22897; IEGG.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C.1.
 DR SMART: SM00044; CLECT; 1.
 DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
 SQ SEQUENCE 404 AA: 45775 MW: A2FA24601453400 CRC64.

```

Query Match          24.0%  Score 78; DB 4; Length 404
Best Local Similarity 100.0%; Pred. No. 4, 4e-72;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONITOLKAAGVELSEKSKIQEIYQETITOLKAAGVELPEKSKLOEIYQETITOLKAA 117
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 74 QDAIYONITOLKAAGVELSEKSKIQEIYQETITOLKAAGVELPEKSKLOEIYQETITOLKAA 113

QY 118 VGELEPEKSKLOEIYQETIT 135
      |||||||||||||||||||
DB 134 VGELEPEKSKLOEIYQETIT 151

RESULT 21
Q95L98 PRELIMINARY; PRT: 268 AA.
AC Q95L98;
DT 01-DEC-2001 (TRIMBLrel. 19, Created)
DT 01-DEC-2001 (TRIMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TRIMBLrel. 20, Last annotation update)
DB2 SDC-SIGNIA type III isoform.
DB3 CD209.
DB4 Homo sapiens (Human).
DB5 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DB6 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
DB7 NCBI_TaxID=9606;
DB8 [1]
DB9 SEQUENCE FROM N A
DB10 PubMed=11337487;
DB11 Mummidi S., Carano G., Lam T., Hoeffler A., Telles V., Begum K.,
DB12 Jinnec F., Ahuja S., Ahuja S.K.
DB13 "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
DB14 specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
DB15 Isoforms: Inter-individual Variation in Expression of DC-SIGN
DB16 Transcripts.";
DB17 J Biol. Chem. 276:33196-33212(2001)
DB18 EMBL: AF042227; AKR1A2.1;
DB19 InterPro: IPR001304; Lectin_C;
DB20 Pfam: PF00059; Lectin_C; 1;
DB21 PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DB22 PROSITE: PS50041; C-TYPE_LECTIN_2; 1;
DB23 SEQUENCE 268 AA; 30427 MW; 113EFD0B57486C3 CRC64.

Query Match          21.2%  Score 69; DB 4; Length 298
Best Local Similarity 100.0%; Pred. No. 6, 5e-63;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QDAIYONITOLKAAGVELSEKSKIQEIYQETITOLKAAGVELPEKSKLOEIYQETITOLKAA 117
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 30 QDAIYONITOLKAAGVELSEKSKIQEIYQETITOLKAAGVELPEKSKLOEIYQETITOLKAA 89

QY 118 VGELEPEKSK 126
      |||||||
DB 90 VGELEPEKSK 98

RESULT 22
Q95L98 PRELIMINARY; PRT: 404 AA
AC Q95L98;
DT 01-DEC-2001 (TRIMBLrel. 19, Created)
DT 01-DEC-2001 (TRIMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRIMBLrel. 20, Last annotation update)
DB2 Dendritic cell-specific ICAM-3 grabbing non-integrin (Fragment).
DB3 DC-SIGN.
DB4 Pan Troglodytes (Chimpanzee).
DB5 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DB6 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
DB7 NCBI_TaxID=9598;
DB8 [1]
DB9 SEQUENCE FROM N A
DB10 Grijftinkbeek T.R.H., Koopman G., Van Nuijnhoven G.C.F., Van Vliet S.J.,
DB11 Van Schijndel A.C., Engeling A., Heeney J.L., Van Kooyk Y.

```

```

Query Match          13.2%  Score 43; DB 6; Length 404;
Best Local Similarity 100.0%; Pred. No. 8, 2e-36;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 TYONTIOLKAAGVELSEKSKIQEIYQETITOLKAAGVELPEKSK 103
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 77 TYONTIOLKAAGVELSEKSKIQEIYQETITOLKAAGVELPEKSK 119

RESULT 23
Q95L98 PRELIMINARY; PRT: 381 AA.
AC Q95L98;
DT 01-DEC-2001 (TRIMBLrel. 19, Created)
DT 01-DEC-2001 (TRIMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TRIMBLrel. 20, Last annotation update)
DB2 Dendritic cell-specific ICAM-3 grabbing nonintegrin.
DB3 Macaca mulatta (Rhesus macaque).
DB4 Fukui Y., Matsuoka, Chordata; Craniata; Vertebrata; Euteleostomi;
DB5 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
DB6 Cercopithecoidea; Macaca.
DB7 NCBI_TaxID=9544;
DB8 [1]
DB9 SEQUENCE FROM N A.
DB10 MEDLINE 2145051; PubMed=11581396;
DB11 Parthaud F., Pohlmann S., Sparwasser T., Kimata M.T., Choi Y.K.,
DB12 Hagarty B.S., Ahmad N., MacLellan T., Edwards T.G., Leslie G.J.,
DB13 Aranson J., Reinhardt J.A., Kimata J.T., Littman D.R., Hoxie J.A.,
DB14 Doms R.W.;
DB15 "Functional and antigenic characterization of human, Rhesus macaque,
DB16 pigtailed macaque, and murine dc-sign.";
DB17 J. Virol. 76:16781-16789(2002);
DB18 EMBL: AF369785; AL014438.1;
DB19 InterPro: IPR001304; Lectin_C;
DB20 Pfam: PF00059; Lectin_C; 1;
DB21 PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DB22 PROSITE: PS50041; C-TYPE_LECTIN_2; 1;
DB23 INTERPRO.
DB24 SEQUENCE 381 AA; 42897 MW; 01FE7B0B42C91D49 CRC64;

Query Match          12.3%  Score 40; DB 6; Length 381;
Best Local Similarity 100.0%; Pred. No. 9, 9e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 ETRIKAAAGVELPEKSKLOEIYQETITOLKAAGVELPEKSK 149
      |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 126 ETRIKAAAGVELPEKSKLOEIYQETITOLKAAGVELPEKSK 165

RESULT 24
Q8S0B2 PRELIMINARY; PRT: 381 AA
AC Q8S0B2;
DT 01-JUN-2002 (TRIMBLrel. 21, Created)
DT 01-JUN-2002 (TRIMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRIMBLrel. 21, Last annotation update)
DB2 Type II membrane protein CD209.
DB3 Macaca mulatta (Rhesus macaque).
DB4 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DB5 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
DB6 Cercopithecoidea; Macaca.

```



```

OX NCB1_taxid=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21819420; PubMed=11818554;
RA Wu L., Bashirou A.A., Martin T.D., Villamide L., Mohlhop E.,
  Chertov A.O., Unutara D., Pope M., Carlington M., Kowalramani V.N.;
  "Rhesus macaque dendritic cells efficiently transmit primate
  lentiviruses independently of DC-SIGN."
  Proc. Natl. Acad. Sci. U.S.A. 99:1568-1574(2002).
RL EMBL: AY040319; AAK74185.1; -.
SF SEQUENCE 381 AA; 42955 MW; 073E7B0B42C91D49 CRC64;

Query Match 12.3%; Score 40; DB 6; Length 381.
Best Local Similarity 100.0%; Pred. No. 9, 9e-33;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 ELTRKAAVGLPEPSKLOEIVQELTRKAAVGLPEPSK 149
DB 126 ELTRKAAVGLPEPSKLOEIVQELTRKAAVGLPEPSK 165
|||||
RESULT 25
QY5JC6 PRELIMINARY; PRT; 381 AA.
AC QY5JC6;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Dendritic cell-specific ICAM-3 grabbing nonintegrin.
OS Macaca nemestrina (pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
  Cercopitheciinae; Macaca.
CX NCB1_taxid=9545;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21465051; PubMed=11581396;
RA Haribaud F., Pohlmann S., Sparwasser T., Kimura M.T., Choi Y.K.,
  Haegarty B.S., Ahmad N., Macfarlan T.G., Leslie G.J.,
  Arriason J., Reinhart T.A., Kimata J.T., Ilftman D.R., Hoxie J.A.,
  Doms R.W.;
  "Functional and antigenic characterization of human, rhesus macaque,
  pig-tailed macaque, and murine dc-sigm."
  J. Virol. 75:10281-10289(2001).
RL EMBL: AF343727; AAL14428.1; -.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
RW Integrin.
SF SEQUENCE 381 AA; 42951 MW; C4F6E2F1D454B74A CRC64;

Query Match 10.5%; Score 34; DB 6; Length 381.
Best Local Similarity 100.0%; Pred. No. 1, 6e-26;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 AAVGELPEPSKLOEIVQELTRKAAVGLPEPSK 126
DB 132 AAVGELPEPSKLOEIVQELTRKAAVGLPEPSK 165
|||||
RESULT 26
QY5J96 PRELIMINARY; PRT; 404 AA.
AC QY5J96;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Dendritic cell-specific ICAM-3 grabbing non-integrin.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;

```

```

OC Cercopitheciinae; Macaca.
OX NCB1_taxid=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=LYMPH NODE;
RA van Schijndel A.C., Engering A., Heeney J.L., van Kooyk Y.,
  "Identification of Rhesus and Chimpanzee DC-SIGN, both a
  gp120 trans-receptors similar to human DC-SIGN."
  Submitted (JUN-2001) to the EMBL/GenBank/TrEMBL databases.
RL EMBL: AF310886; AAK97459.1; -.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
RW Integrin.
SF SEQUENCE 404 AA; 45655 MW; F7755A13430CC17 CRC64;

Query Match 8.0%; Score 26; DB 6; Length 404.
Best Local Similarity 100.0%; Pred. No. 3, 4e-18;
Matches 26; Conservative 0; Mismatches 0; Indels 0;

QY 110 ELTRKAAVGLPEPSKLOEIVQEL 145
DB 126 ELTRKAAVGLPEPSKLOEIVQEL 161
|||||
RESULT 27
QY6Q06 PRELIMINARY; PRT; 168 AA.
AC QY6Q06;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE MDC-SIGNA type IV isoform.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Hominidae;
  Homininae; Homo.
CX NCB1_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11337487;
RA Mummidi S., Catano G., Lam L., Hoxie A., Telles V., Bee
  Jimenez F., Ahuja S.S., Ahuja S.K.;
  "Extensive repertoire of Membrane-bound and Soluble Integrin
  specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and its
  isoforms. Inter-individual Variation in Expression of DC
  Transcripts."
  J. Biol. Chem. 276:33196-33212(2001).
RL EMBL: AY042224; AAK91849.1; -.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
SF SEQUENCE 168 AA; 18646 MW; 9017187C9E489F59C CRC64;

Query Match 6.5%; Score 21; DB 4; Length 168.
Best Local Similarity 100.0%; Pred. No. 2, 4e-13;
Matches 21; Conservative 0; Mismatches 0; Indels 0;

QY 253 WNGSLDNLQEGTQWQVDSPL 273
DB 79 WNGSLDNLQEGTQWQVDSPL 99
|||||
RESULT 28
QYV1K4 PRELIMINARY; PRT; 207 AA.
AC QYV1K4;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)

```

DE SIGNR3.
 GN CD2099 OR SIGNR3.
 OS Mus musculus (Mouse).
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
 CC Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
 OX NCBI_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RF Park C.G., Steinman R.M.;
 RI "Alternatively Spliced Forms of Mouse DC-SIGN Homologs."
 RI Submitted (Oct-2001) to the FMBI/Genbank/DBJ databases.
 DE FMBI: AF440280; AAL33584.1;
 DE MGI: MGI:2157947; CD209d.
 DE InterPro: IPR001304; Lectin_C.
 DE Pfam: PF00059; Lectin_C; 1.
 DE SMART: SM00034; CLECT; 1.
 DE PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DE PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 SC SEQUENCE 207 AA, 23500 MW, 16FMHLL47FCAC3FR CCK64.

Query Match 3.4%; Score 11; DB 11; Length 207;
 Best Local Similarity 100.0%; Pred No. 0.0065;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 WDCSPLSPSF 277
 DB 148 WDCSPLSPSF 158
 |||||

RESULT 29
 Q91ZM8 PRELIMINARY; PRT; 237 AA
 ID Q91ZM8
 AC Q91ZM8.
 DT 01-DEC-2001 (TEMBLrel. 19, Created)
 DT 01-DEC-2001 (TEMBLrel. 21, Last annotation update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE SIGNR3.
 GN CD209d OR SIGNR3.
 OS Mus musculus (Mouse).
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
 CC Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
 OX NCBI_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RF PubMed=11581173;
 RF Park C.G., Takahara K., Umemoto E., Yashima Y., Matsubara K.,
 RA Matsuda Y., Clausen B.E., Inaba K., Steinman R.M.;
 RI "Five mouse homologues of the human dendritic cell C-type lectin, DC-SIGN."
 RI Submitted (Jan-2001) to the FMBI/Genbank/DBJ databases.
 DE FMBI: AF373411; AAL1237.1;
 DE MGI: MGI:2157947; CD209d.
 DE InterPro: IPR001304; Lectin_C.
 DE Pfam: PF00059; Lectin_C; 1.
 DE PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DE PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 SC SEQUENCE 237 AA, 26925 MW, 52EEKQKCEAF8HFR CCK64;

Query Match 3.4%; Score 11; DB 11; Length 217;
 Best Local Similarity 100.0%; Pred No. 0.0073;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 267 WDCSPLSPSF 277
 DB 178 WDCSPLSPSF 188
 |||||

RESULT 30
 Q16641 PRELIMINARY; PRT; 168 AA
 AC Q16641
 O16641

DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01 JAN 1998 (TEMBLrel. 05, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE A hypothetical 18.2 kDa protein.
 DE W09G10.5.
 GN Caenorhabditis elegans.
 OC Eukaryota, Metazoa, Nemata, Chordata, Craniata, Vertebrata, Euteleostomi,
 CC Rabditidae, Polychordata, Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE: 99059613; PubMed: 9851916;
 RA Neece;
 RI "Genome sequence of the nematode C. elegans: a platform for
 RI Investigating biology. The C. elegans Sequencing Consortium."
 RI Science 282:2012-2018(1998).
 RN 121
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Bentley D., Goebl D., Holmes A.;
 RI "The sequence of C. elegans cosmid W09G10."
 RI Submitted (Aug 1997) to the FMBI/Genbank/DBJ databases.
 RN 131
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RA Waterston R.;
 RI "Direct Submission."
 RI Submitted (SEP-2001) to the FMBI/Genbank/DBJ databases.
 DR EMBL: AF016671; AAF66113.1;
 DR HSSP: P22897; IFCG.
 DR InterPro: IPR001304; Lectin_C.
 DE Pfam: PF00059; Lectin_C; 1.
 DE SMART: SM00034; CLECT; 1.
 DE PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 DR Hypothetical protein.
 SC SEQUENCE 108 AA, 18217 MW, 494FE3ADU031BRC5 CCK64;

Query Match 2.8%; Score 9; DB 5; Length 168;
 Best Local Similarity 100.0%; Pred No. 0.64;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 YWNSCEPNN 288
 DB 123 YWNSCEPNN 131
 |||||

RESULT 31
 Q91ZX1 PRELIMINARY; PRT; 238 AA
 ID Q91ZX1
 AC Q91ZX1.
 DT 01 DEC 2001 (TEMBLrel. 19, Created)
 DT 01 DEC 2001 (TEMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE DC-SIGN (Dendritic cell C-type lectin protein CIRE).
 GN CD209A OR DCSIGN.
 OS Mus musculus (Mouse).
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
 CC Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
 OX NCBI_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RF PubMed=11581173;
 RF Park C.G., Takahara K., Umemoto E., Yashima Y., Matsubara K.,
 RA Matsuda Y., Clausen B.E., Inaba K., Steinman R.M.;
 RI "Five mouse homologues of the human dendritic cell C-type lectin, DC-SIGN."
 RI Submitted (Jan-2001) to the FMBI/Genbank/DBJ databases.
 DE FMBI: AF373411; AAL1237.1;
 DE MGI: MGI:2157947; CD209a.
 DE InterPro: IPR001304; Lectin_C.
 DE Pfam: PF00059; Lectin_C; 1.
 DE PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DE PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 SC SEQUENCE 238 AA, 26925 MW, 52EEKQKCEAF8HFR CCK64;

RA Brodnicki T.C., Lew A.M., Shortman K., Wright M.D.:
 RT "Molecular cloning of a C-type lectin superfamily protein
 RT differentially expressed by CD4⁺CD45⁺ splenic dendritic cells."
 RL Mol. Immunol. 38:365-373(2001).
 DR EMBL: AF373408; AAL13234.1; .
 DR EMBL: AY049062; AAK85825.1; .
 DR MGI: 2157943; Cd209a.
 DR InterPro: IPR001304; Lectin_C.
 DR Pfam: PF00059; Lectin_C_1.
 DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNWN_1.
 DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
 KW Lectin.
 SQ SEQUENCE 238 AA: 27149 MW: 05025PF02564DC2 CRC64:
 Query Match 2.88; Score 9; DB 1; Length 238;
 Best Local Similarity 100.0%; Pred. No. 0.87;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 YQELTOLKA 93
 |||||||
 DB 91 YQELTOLKA 99

RESULT 32
 Q8TB07 PRELIMINARY: PRT: 491 AA.
 ID Q8TB07
 AC Q8TB07
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DR Hypothetical 53.0 kDa protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RX SEQUENCE FROM N.A.
 RP STRAUSBERG R.;
 RC TISSUE=UTERUS;
 RA Strausberg R.;
 RL Submitted (FEF8-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC024214; AAI24214.1; .
 KW Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 491 AA: 53045 MW: BAC0138C38496A08 CRC64:
 Query Match 2.88; Score 9; DB 4; Length 491;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 AGVLAIV 42
 |||||||
 DB 278 AGVLAIV 286

RESULT 33
 Q96002 PRELIMINARY: PRT: 34 AA.
 ID Q96002
 AC Q96002
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DR SDC-STIGNIA type IV isoform.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Granulata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RX SEQUENCE FROM N.A.
 RP PubMed:113374877;
 RA Jimenez F., Cabana G., Lam L., Hocile A., Telles V., Begum K.,
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;
 RT "Extensive Repertoire of Membrane-Bound and Soluble Dendritic Cell-

RT Specific 10AM-3-grabbing Neofectin 1 (ncf-3n1) and its
 RT isoforms. Inter-individual variation in expression of the
 RT transcripts."
 RL J. Biol. Chem. 276:33196-33212(2001).
 DR F044: AY042228; AAK91853.1; .
 SQ SEQUENCE 34 AA: 4046 MW: 51FA66911DAA150 CRC64:
 Query Match 2.59; Score 8; DB 4; Length 34;
 Best Local Similarity 100.0%; Pred. No. 1.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 1 MSDSKEPR 8
 |||||||
 DB 1 MSDSKEPR 8

RESULT 34
 Q9LFZ0 PRELIMINARY: PRT: 88 AA.
 ID Q9LFZ0
 AC Q9LFZ0
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE T7N9.1
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tr
 OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudi
 OC eustids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RX SEQUENCE FROM N.A.
 RA Khan S., Brooks S., Buchler E., Chao G., Johnson-Hopson
 RA Shion P., Altali H., Bel G., Chin G., Chou J., Choi E.
 RA Conway A., Gonzales A., Hansen N., Howard B., Koo T., Lam
 RA Lenz G., Li J., Liu A., Liu K., Liu S., Makarewicz N., M
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Toriumi M., Vaysberg M., Yu G., Friedtschel N.A., Theodor
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC T7N9.1
 RT T."
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases
 RN 12
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases
 RN 13
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases
 RN 14
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases
 RN 15
 RP SEQUENCE FROM N.A.
 RA Cheuk P., Shinn P., Brooks S., Buchler E., Chao G., J
 RA Khan S., Kim C., Altali H., Bel G., Chin G., Chou J.
 RA Conn J., Conway A., Gonzales A., Hansen N., Howard B.,
 RA Lee J., Lenz G., Li J., Liu A., Liu J., Liu S., Makarew
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., S
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R.,
 RA Theologis A., Ecker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases
 RN 16
 RP SEQUENCE FROM N.A.
 RA Cheuk P., Shinn P., Brooks S., Buchler E., Chao G., J
 RA Khan S., Kim C., Altali H., Bel G., Chin G., Chou J.
 RA Conn J., Conway A., Gonzales A., Hansen N., Howard B.,
 RA Lee J., Lenz G., Li J., Liu A., Liu J., Liu S., Makarew
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., S
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R.,
 RA Theologis A., Ecker J.;
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases

D> EMBL: AC000348; AAF79848.1; -
 S> SEQUENCE: 88 AA; 10203 MW; 494d0149a1809444 CRC64;
 Query Match: 2.5%; Score 8, DB 10, Length 88;
 Best Local Similarity 100.0%; Pred. No. 3.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Q> 25 L0LSEFM.32
 Db 38 L0LSEFM.45
 RESULT 35
 Q91ZW7 PRELIMINARY: PPT: 298 AA.
 ID Q91ZW7;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DI 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DF SIGNR4.
 GN C02094 OR SIGNR4.
 OS Mus musculus (Mouse)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 PN 11
 RE SEQUENCE FROM N A
 RC STRAIN=C57BL/6;
 RX PubMed-11581173;
 RA Park G.G., Takahara K., Umecoro E., Yashima Y., Matsubara K.,
 RA Matsuda Y., Clausen R.E., Inaba K., Stolman R.M.;
 RT "Five mouse homologues of the human dendritic cell C-type lectin, DC-
 RT SIGN-1".
 RL Int. Immunol. 13:1283-1290(2001).
 DR EMBL: AF373412; AAL13238.1; -
 DR MDP: MG1:2157948; C02096.
 DR InterPro: IPR001304; Lectin_C
 DR Pfam: PF00059; Lectin_C_1
 DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DF PROSITE: PS00615; C-TYPE_LECTIN_2; 1.
 SE SEQUENCE: 208 AA; 24382 MW; 219cf71eb870eb7 CRC64.
 Query Match: 2.5%; Score 8, DB 11, Length 208;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Q> 253 WMGLSDLN 260
 Db 135 WMGLSDLN 142
 RESULT 36
 Q94ZD4 PRELIMINARY: PPT: 279 AA.
 ID Q94ZD4;
 AC Q94ZD4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DI 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE A1912410/F5011_7.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 GN NBI_TaxID=3702;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Cheuk P., Chen H., Kim C.T., Kossuma F., Meyers M.C., Barth T.,
 RA Bowser L., Carlini P., Dale J.M., Goldsmith A.D., Hashimoto Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Natusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis P.W., Theologis A.,

RA Ecker J.R.;
 RT "Arabidopsis cDNA clones.";
 RL Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY057617; AAL44112.1; -
 DR MEROPS: S14.001; -
 DF InterPro: IPR001907; CLP_protease.
 DF Pfam: PF00574; CLP_protease_1.
 DR TrEMBL: T1G00493; CLP1.1.
 SE SEQUENCE: 279 AA; 31150 MW; F7FFD2C3374830B CRC64.
 Query Match: 2.5%; Score 8; DB 10; Length 279;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Q> 262 ICTGQWD 269
 Db 82 ICTGQWD 89
 RESULT 37
 Q9XJ36 PRELIMINARY: PPT: 279 AA.
 ID Q9XJ36;
 AC Q9XJ36;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DI 01-NOV-1999 (TrEMBLrel. 12, last sequence update)
 DE 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE NCLPP3 (F5011.13) (Similar to nclpp3 db|BA082066.1).
 GN NCLPP3 OR A1912410, F5011_13 OR A1912410.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 GN NBI_TaxID=3702;
 RN 11
 RP SEQUENCE FROM N.A.
 RA STRAIN=COLOMBIA;
 RA Navechayashvili K., Tre M., Kiyosue T., Shinzaki K., Watanabe A.;
 RT "Identification of clp genes expressed in senescing Arabidopsis
 RT leaves.";
 RL Plant Cell Physiol. 0:0-0(1999).
 DE 13
 RP SEQUENCE FROM N.A.
 RA Chao Y., Brooks S., Buchler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shino P., Altati H., Bel G., Chin C., Chou J., Choi E., Conn L.,
 RA Conway A., Gonzalez A., Hansen R., Howing B., Koo T., Lam B., Lee J.,
 RA Lenz G., Li J., Liu A., Liu K., Liu S., Mukharasy N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vayenberg M., Yu G., Federspiel N.A., Theologis A.,
 RA Ecker J.R.;
 RT "Genomic sequence for Arabidopsis thaliana BAC F5011 from chromosome
 RT 1.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN 13
 RP SEQUENCE FROM N.A.
 RA Ecker J.R.;
 RN Submitted (MAR 2000) to the EMBL/Genbank/DBJ databases.
 RT Ecker J.R.;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 RN 15
 RP SEQUENCE FROM N.A.
 RA Cheuk P., Shino P., Brooks S., Buchler E., Chao G., Johnson-Hopson C.,
 RA Khan S., Kim C., Altati H., Bel G., Chin C., Chou J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen R., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz G., Li J., Liu A., Liu J., Liu S., Mukharasy N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vayenberg M., Yu G., Davis P., Federspiel N.,
 RA Theologis A., Ecker J.;
 PL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases
 RN 16
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,

```

RA Palm C.J., Bowser L., Jones T., Bant J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shin P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam H., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Bant J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shin P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB022327; BAA82066.1;
DR EMBL: AC025416; AAF79635.1;
DR EMBL: AY062770; AAL32848.1;
DR EMBL: AY081641; AAM10203.1;
DR HSSP: P19245; ITF;
DR MEMOS: S14.001;
DR InterPro: IPR001907; CLP-Protease;
DR Pfam: PF00574; CLP-Protease;
DR PRINTS: PR00127; CLP-PROTEASEP;
DR Trifams: TIGR00494; CLP: 1
SU SEQUENCE 279 AA; 31208 MW; DADICFDDN478408 CR664;

Query Match 2.58; Score 8; DB 10; Length 279;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 ECTMOWND 269
DB H2 ECTMOWND 89

RESULT 38
ID P72781 PRELIMINARY; PRT: 282 AA.
AC P72781;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Regulatory components of sensory transduction system.
GN SLR1783.
OS Synecchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecchocystis
CX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97061201; Pubmed 8905231.
RA Kaneke T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Saitou M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Jimra S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Taira S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecchocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-146(1996)
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER
CC REGULATORY COMPONENTS OF SENSOR TRANSDUCTION SYSTEMS.
CC -1- SIMILARITY: BELONGS TO THE LUXR/DHAP FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL: D90900; BAA16796.1;
DR HSSP: P08402; 1B00.
DR InterPro: IPR000792; HTH_LuxR.
DR InterPro: IPR001789; Response_Treg.
DR Pfam: PF00196; Gereg_1.
DR Pfam: PF00072; response_Treg; 1.
DR Prodom: PD000039; response_Treg; 1.
DR Prodom: PD000307; HTH_LuxR; 1.
DR SMART: SM00421; HTH_LuxR; 1.

```

```

DR SMART: SM00448; REC; 1.
KW DNA binding; Phosphorylation; Sensory transduction;
KW Transcription regulation; Complete proteins.
SU SEQUENCE 282 AA; 31395 MW; IAT25612H766B457 CR674;

Query Match 2.58; Score 8; DB 16; Length 28;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 149 KLOEYOE 156
DB PRO ECTEYOE 156

RESULT 39
ID Q9DBV4 PRELIMINARY; PRT: 41 AA.
AC Q9DBV4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE 1810030122RIK protein.
CN CD2098 OR 1810030122RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Cladoda; Vertebrata; Eute-
OC Mammalia; Eutheria; Rodentia; Sciurodonta; Muridae; Mur
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=PANCREAS;
RA MEDLINE=21085660; Pubmed=1121781;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yae
RA Saito T., Okazaki Y., Gojohri I., Hone H., Kasukawa Y.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casady
RA Fleischmann W., Gaasterland T., Gissi C., King R., Koehle
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G.,
RA Schindl L.M., Staudt F., Suzuki K., Tamita M., Wagner J.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R.,
RA Blake J., Boffelli D., Hojima N., Carninci P., de Heide
RA Brownstein M.J., Bult C., Fletcher G., Fujita M., Garib
RA Gustincich S., Hill D., Holtmann M., Hume D.A., Kamiya M.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mont
RA Nordone P., Rind B., Ringwald M., Rodriguez L., Sakamoto
RA Sasaki H., Sato K., Schoenbach G., Seya T., Shibata Y.,
RA Suzuki H., Taya-Oka K., Wani K., Weller G., Whitaker
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai J., K
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collec-
RL Nature 409:985-990(2001).
DR EMBL: AC087056; BAA25105.1;
DR HSSP: P22897; 1P6G.
DR MED: M01191615; CD209B.
DR InterPro: IPR001353; AntiProtein_1.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C; 1.
DR PRINTS: PR00356; ANTIPEPTEIN.
DR SMART: SM00034; CLCTP; 1.
DR PROSITE: PS00615; G_TYPE_LETIN_1; UNKNOWN_1.
DR PROSITE: PS50041; G_TYPE_LETIN_2; 1.
SU SEQUENCE 311 AA; 35618 MW; 92355030ZEF9041 CR674;

Query Match 2.58; Score 8; DB 11; Length 30;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY 23 LVLOLSE 30
DB 40 LVLOLSE 47

RESULT 40

```

```

091ZX0
ID 091ZX0 PRELIMINARY: PRT: 325 AA.
AC 091ZX0:
DT 01-DEC-2001 (TREMblrel, 19, Created)
DT 01-DEC-2001 (TREMblrel, 19, Last sequence update)
DE 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE SIGNR1.
GN CD209H OR SIGNR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 11
RC SEQUENCE FROM N A
RC STRAIN=C57BL/6;
RX PubMed-11581173;
KA Park C.G., Takahara K., Imamoto F., Yashima Y., Matsubara K.,
KA Matsuda Y., Clausen B.F., Inaba K., Steinman R.M.;
R1 "Five mouse homologues of the human dendritic cell C-type lectin, DC
R1 SIGN.";
R1 Int. Immunol. 13:1283-1290(2001).
DR EMBL: AF373409; AAL13235.1; -.
DR MGD: MGI:1916415; G62998.
DR InterPro: IPR001304; Lectin_C_1.
DR Pfam: PF00059; Lectin_C_1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE: PS00641; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 325 AA; 37111 MW; 9C938E40/247C64 CRC64;

Query Match 2.58; Score 8; DB 11; Length 325;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 23 LVGLLSE 30
|11111111
DB 54 LVGLLSE 61

RESULT 41
054910
ID 054910 PRELIMINARY: PRT: 355 AA.
AC 054910:
DT 01-NOV-1996 (TREMblrel, 01, Created)
DT 01-NOV-1996 (TREMblrel, 01, Last sequence update)
DE 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Oligopeptidase.
GN OPPD.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN 11
RC SEQUENCE FROM N.A.
RC STRAIN=SEKOTYPE M49;
RX MEDLINE=87039725; PubMed-8885277;
KA Podbielski A., Pohl B., Moisschnik M., Koerner C., Schmidt K.H.,
KA Rodzinski E., Leonard R.A.B.;
R1 "Molecular characterization of a group A streptococcal (GAS)
R1 Oligopeptidase (Opp) and its effect on cysteine protease
R1 production.";
R1 Mol. Microbiol. 21:1087-1099(1996)
DR EMBL: X89237; CAA61525.1; -.
DR InterPro: IPR003593; AAA_Atpase
DR InterPro: IPR003439; ABC_Transport;
DR Pfam: PF00005; ABC_tran; 1.
DR PRODOM: PD000006; ABC_Transport; 1.
DR SMART: SM00382; AAA; 1.
DR ATP-binding; Complete proteome.
SQ SEQUENCE 356 AA; 39645 MW; E567C7F94C52556A CRC64;

Query Match 2.58; Score 8; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 23 LVGLLSE 30
|11111111
DB 54 LVGLLSE 61

```

```

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 178 QELTUKT 185
|11111111
DB 78 QELTUKT 85

RESULT 42
09A1F7
ID 09A1F7 PRELIMINARY: PRT: 356 AA.
AC 09A1F7:
DT 01-JUN-2001 (TREMblrel, 17, Created)
DT 01-JUN-2001 (TREMblrel, 17, Last sequence update)
DE 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Oligopeptidase (ATP-binding protein).
GN OPPD OR SPY0296.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN 11
RC SEQUENCE FROM N.A.
RC STRAIN=SEKOTYPE M49;
RX MEDLINE=87039725; PubMed-8885277;
KA Podbielski A., Pohl B., Moisschnik M., Koerner C., Schmidt K.H.,
KA Rodzinski E., Leonard R.A.B.;
R1 "Molecular characterization of a group A streptococcal (GAS)
R1 Oligopeptidase (Opp) and its effect on cysteine protease
R1 production.";
R1 Mol. Microbiol. 21:1087-1099(1996)
DR EMBL: X89237; CAA61525.1; -.
DR InterPro: IPR003593; AAA_Atpase
DR InterPro: IPR003439; ABC_Transport;
DR Pfam: PF00005; ABC_tran; 1.
DR PRODOM: PD000006; ABC_Transport; 1.
DR SMART: SM00382; AAA; 1.
DR ATP-binding; Complete proteome.
SQ SEQUENCE 356 AA; 39703 MW; E565AEFP4C52556A CRC64;

Query Match 2.58; Score 8; DB 10; Length 356;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

UY 178 QELTUKT 185
|11111111
DB 78 QELTUKT 85

RESULT 43
082594
ID 082594 PRELIMINARY: PRT: 382 AA.
AC 082594:
DT 01-NOV-1998 (TREMblrel, 08, Created)
DT 01-NOV-1998 (TREMblrel, 08, Last sequence update)
DE 01-DEC-2001 (TREMblrel, 19, Last annotation update)
DE FliC4.2 protein.
GN FliC4.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Eudicotyledons; Core eudicot; Rosidae;
OC Eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN 11
RC SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX Washin;
R1 "The A. thaliana Genome Sequencing Project.";
R1 Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.
RN 12
RC SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX Abu-Threideh D., Steneking-Lis, Larsson Y., Trevisan E.;

```

```

KT      "the sequence of A. thaliana F1104."
RL      Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CV. COLUMBIA;
RA      Waterston R.;
RL      Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF096470; AAC62779.1;
DR      InterPro: IPR000477; PVT5;
DR      InterPro: IPR000531; TonB_boxC;
DR      Pfam: PF00078; rvt; 1;
DR      PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW      RNA-directed DNA polymerase
SQ      SEQUENCE 382 AA: 43977 MW: 0DE2342C21D03745 CRC64;

Query Match          2.5%: Score 8; DB 10; Length 382;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      47 VPSSLQSE 54
DB      165 VPSSLQSE 172

RESULT 44
Q9C1R9          PRELIMINARY:      PRT:      441 AA.
ID      Q9C1R9
AC      Q9C1R9;
DT      01-JUN-2001 (TREMblrel. 17, Created)
DT      01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT      01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE      Centromere binding protein 1 (Centromere binding factor 1)
GN      CBF1.
OS      Candida glabrata (Yeast) (Torulopsis glabrata)
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX      NCBI_TaxID=5478;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21332095; PubMed=11438645;
RA      Stoyan T., Gloeckner G., Diekmann S., Carbon J.;
RT      "Multifunctional Centromere Binding Factor 1 Is Essential for
RT      Chromosome Segregation in the Human Pathogenic Yeast Candida
RT      glabrata."
RL      Mol. Cell. Biol. 21:4875-4888(2001).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Stoyan T., Gloeckner G., Diekmann S., Carbon J.;
RT      "Cloning of a centromere binding factor (CBF1) gene from Candida
RT      glabrata."
RL      Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AF233433; AAK14999.1;
DR      EMBL: AF159250; AAK58891.1;
DR      HSSP: P22415; IAN4.
DR      InterPro: IPR001092; H1H_Basic;
DR      Pfam: PF00010; H1H_1;
DR      PRINTS: PR01632; POWDCALPHAI;
DR      SMART: SM00353; HCH; 1;
SQ      SEQUENCE 441 AA: 51409 MW: 477955CD32F9BCB0 CRC64;

Query Match          2.5%: Score 8; DB 3; Length 441;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      154 YQELTRLK 161
DB      367 YQELTRLK 374

RESULT 45
Q9TVX8          PRELIMINARY:      PRT:      472 AA.
AC      Q9TVX8; 017975;

```

```

DT      01-MAY-2000 (TREMblrel. 13, Created)
DT      01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      R31.3 protein (OSM-6 protein).
GN      R31.3 OR OSM-6.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; R;
OC      Rhabditidae; Telodoridae; Caenorhabditis.
OX      NCBI_TaxID=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Harris B.R.;
PI      Submitted (MAY 1996) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=9909613; PubMed=9851916;
RA      none;
RT      "Genome sequence of the nematode C. elegans: A platform for
RT      investigating biology."
RL      Science 287:2012-2018(1998).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=BRISTOL N2;
RA      Collet J., Spike C.A., Lundquist E.A., Shaw J.E., Hernan
RT      "Analysis of osm-6, a gene that affects sensory cilium s
RT      sensory neuron function in Caenorhabditis elegans."
RL      Genetics 148:187-200(1998).
DR      EMBL: Z75954; CAB61020.2; JOURNAL;
DR      EMBL: Z75956; CAB61020.2; JOURNAL;
DR      EMBL: Z75954; CAB61020.2; JOURNAL;
DR      EMBL: AJ000259; CAB03975.1;
SQ      SEQUENCE 472 AA: 53154 MW: 5FA6924EB164278F CRC64;

Query Match          2.5%: Score 8; DB 5; Length 472;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0;

QY      169 DQSKQQU1 176
DB      21 DQSKQQU1 28

```

Search completed: December 7, 2002, 11:16:57
Job time : 39 secs

The first part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present. The second part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present. The third part of the paper discusses the importance of the study of the history of the United States. It is argued that the study of history is essential for a full understanding of the present.